Université de Nantes

UFR de Psychologie

****

ETUDE LONGITUDINALE PREDICTIVE DE LA REUSSITE OU DE L'ECHEC DANS L'APPRENTISSAGE DE LA LECTURE EN LANGUE ARABE

THESE

POUR L'OBTENTION DU GRADE DE DOCTEUR DE L'UNIVERSITE DE NANTES

DISCIPLINE : PSYCHOLOGIE

Présentée et soutenue publiquement par :

Fatma SNOUSSI BOUZOUAYA
Le 09 novembre 2009

Thèse Dirigée par Agnès Florin, Professeur

Co-encadrée par Philippe Guimard, Maître de Conférences

****

Jury :

Riadh BEN REJEB	Professeur, Université de Tunis
Séverine CASALIS	Professeur, Université de Lille 3
Jean ECALLE	Professeur, Université de Lyon 2
Agnès FLORIN	Professeur, Université de Nantes
Philippe GUIMARD	Maître de conférences, Université de Nantes

Remerciements

Je tiens à adresser mes plus vifs et sincères remerciements à tous ceux et celles qui, par leur aide ou par leur présence, ont contribué à la réalisation de ce travail.

Je tiens plus particulièrement à remercier ma directrice de thèse, le Pr. Agnès Florin qui, par sa disponibilité et ses conseils avisés m'a apporté le soutien nécessaire à la réalisation de ce travail dans de bonnes conditions. Qu'elle trouve ici toute ma gratitude.

Je remercie très sincèrement mon co-encadrant, Mr. Philippe Guimard pour ses précieux conseils et ses qualités humaines.

Je voudrais remercier les membres de jury pour avoir accepté de juger mon travail de thèse et pour le temps qu'ils ont consacré à la lecture de ce travail.

Je tiens également à remercier Mme Michèle Belajouza pour l'intérêt qu'elle a toujours porté à mes travaux de recherche. La confiance qu'elle m'a témoignée fut un précieux stimulant.

Un grand merci aux membres du Laboratoire « Education, Cognition et Développement » (Labécd EA 3259) de l'Université de Nantes pour leur accueil et leur soutien.

J'adresse aussi mes remerciements à tous les enseignants du département de psychologie de l'Institut Supérieur des Sciences Humaines de Tunis.

Toute ma reconnaissance s'adresse également à mes proches et mes amis qui depuis tant d'années m'encouragent et dont la chaleur, l'enthousiasme et le soutien m'ont permis de garder le moral :

A mon mari Khaled, je le remercie pour son affectueux soutien et sa courageuse patience. Sans sa présence, son aide et ses prouesses informatiques, rien n'aurait été possible.

A mes enfants, Youssef et Mehdi, avec toute mon affection.

A toute ma famille, et notamment mes parents Mohamed Hédi et Naziha, ma sœur Elhem et son mari Mondher et la petite Mayssa ; Soumaya et Aroua et mon frère Oussama et ma belle famille, Mamie, Héla et Ahmed, Kaouther et Samy, Mohamed et Jihène sans oublier la petite Salima ; qu'ils trouvent ici l'expression de ma reconnaissance pour leur aide et leurs affectueux encouragements.

A tous mes amis et surtout à Nahed, Amal, Anne-Sophie et Aida.

Enfin, je suis très reconnaissante vis-à-vis des directeurs d'écoles, des enseignants et surtout des élèves qui se sont prêtés avec joie au travail expérimental.

Vous avez chacun à votre manière permis l'aboutissement de ce travail et je vous en serais éternellement reconnaissante.

Sommaire

AVANT-PROPOS 8

PREMIERE PARTIE : CADRE THEORIQUE DE LA RECHERCHE 13

CHAPITRE 1 : L'APPRENTISSAGE DE LA LECTURE 14

1. L'IDENTIFICATION DES MOTS ECRITS 14

1.1. Définition 14

1.2. L'identification des mots écrits chez l'apprenti lecteur 16

1.2.1. La médiation phonologique 16

1.2.2. Rôle du contexte dans l'identification des mots 18

1.3. L'identification de mots écrits en lecture experte 20

1.3.1. L'assemblage phonologique chez le lecteur expert 21

1.3.2. L'importance de l'automatisation des procédures d'identification des mots 23

2. LA COMPREHENSION EN LECTURE 25

2.1. Les mécanismes cognitifs de la compréhension 25

2.2. Facteurs de difficulté en compréhension 27

3. LES PRINCIPAUX MODELES DE L'APPRENTISSAGE DE LA LECTURE 28

3.1. Les modèles développementaux en stades 29

3.2. Les modèles de lecture par analogie 31

3.3. Les modèles connexionnistes et interactifs 34

3.3.1. Le modèle connexionniste de Seidenberg & Mc Clelland (1989) 34

3.3.2. Le modèle de Gombert, Bryant & Warrick (1997) 37

4. SPECIFICITES DE LA LECTURE EN LANGUES SEMITIQUES 39

5. CONCLUSION 42

CHAPITRE 2 : LA LANGUE ARABE : SPECIFICITES ET APPRENTISSAGE 44

1. HISTORIQUE 44

2. CARACTERISTIQUES DE LA LANGUE ARABE 45

2.1. Système d'écriture 45

2.2. Système morphologique 48

3. LA SITUATION DE DIGLOSSIE 51

4. SPECIFICITES DE L'APPRENTISSAGE DE LA LECTURE EN LANGUE ARABE 56

4.1. L'évolution des stratégies de lecture en arabe 56

4.2. Rôle des voyelles et du contexte 59

5. CONCLUSION 62

CHAPITRE 3 : LES DETERMINANTS DE LA REUSSITE DANS L'APPRENTISSAGE DE LA LECTURE 64

1. LES CONNAISSANCES PHONOLOGIQUES 64

1.1. Les unités de traitement phonologique 65

1.2. Développement des capacités métaphonologiques 69

1.3. Apprentissage de la lecture et conscience phonologique : quels liens ? 73

2. LES CONNAISSANCES MORPHOLOGIQUES 77

2.1. La place de la morphologie dans l'apprentissage de la lecture 78

2.2. Rôle des connaissances morphologiques dans l'apprentissage de la lecture 82

3. LA CONNAISSANCE DES LETTRES 87

4. LA DENOMINATION RAPIDE 90

5. COMPETENCES COGNITIVES ET APPRENTISSAGE DE LA LECTURE 91

5.1. La mémoire de travail 92

5.2. Rôle des fonctions exécutives 95

5.3. Le niveau intellectuel 96

5.4. Les capacités visuo-attentionnelles 98

6. INFLUENCE DES METHODES D'ENSEIGNEMENT 100

7. EFFETS DES CARACTERISTIQUES LINGUISTIQUES 104

8. LE POIDS DES FACTEURS SOCIAUX 108

8.1. La représentation sociale sur la lecture 108

8.2. Les pratiques et les conceptions éducatives des parents 109

9. CONCLUSION 110

CHAPITRE 4:PREDICTEURS DE L'APPRENTISSAGE DE LA LECTURE 113

1. QUELLES HABILETES PRECOCES PREDISENT LES COMPETENCES ULTERIEURES DE LECTURE ? 113

1.1. Etudes longitudinales prédictives 114

1.2. Etudes comparatives interlangues 119

1.3. Etudes comparatives entre bons lecteurs et faibles lecteurs 122

1.4. Etudes avec entraînement métalinguistique 125

2. PREDICTION DE LA REUSSITE SCOLAIRE FOURNIE PAR LES ENSEIGNANTS 128

3. PREDICTIONS INDIVIDUELLES 131

4. APPORTS ET LIMITES DES ETUDES PREDICTIVES 132

5. CONCLUSION 134

DEUXIEME PARTIE : PARTIE EXPERIMENTALE 136

CHAPITRE 1: PERSPECTIVES DE LA RECHERCHE 137

1. PRESENTATION DE LA PROBLEMATIQUE 137
2. OBJECTIFS ET HYPOTHESES DE LA RECHERCHE 143
3. TRAITEMENT STATISTIQUE DES DONNEES 148
CHAPITRE 2 : COMPETENCES EVALUEES EN CLASSE PREPARATOIRE ET PREDICTION DU NIVEAU ULTERIEUR EN LECTURE 151
1. METHODE 151
1.1. Sujets 151
1.2. Matériel 152
1.2.1. Mesures proposées en classe préparatoire (T1) 152
1.2.1.1. Evaluation du langage oral 152
1.2.1.2. Evaluation du langage écrit 157
1.2.1.3. Evaluation des compétences cognitives 157
1.2.2. Mesures proposées en 1ère année primaire (T2) 159
1.2.3. Mesures proposés en 2ème année primaire (T3) 160
1.3. Procédure 161
2. RESULTATS 162
2.1. Analyse des performances à T1, T2 et T3 162
2.1.1. Performances aux épreuves langagières en classe préparatoire (T1) 162
2.1.2. Performances aux épreuves cognitives en classe préparatoire (T1) 164
2.1.3. Performances aux épreuves langagières en 1ère année (T2) 164
2.1.4. Performances aux épreuves de lecture à T2 et T3 166
2.1.4.1. Les performances en identification de mots à T2 166
2.1.4.2. Les performances en identification de mots à T3 166
2.1.4.3. Les performances en lecture-compréhension à T3 168
2.2. Analyse du lien entre les scores aux différentes épreuves 168
2.2.1. Analyse des corrélations 168
2.2.1.1. Corrélations entre les épreuves en classe préparatoire (T1) 168
2.2.1.2. Corrélations entre les épreuves en 1ère année (T2) 170
2.2.1.3. Corrélations entre les épreuves à T1 et T2 170
2.2.1.4. Corrélation entre les épreuves de lecture à T2 et T3 171
2.2.2. Analyse factorielle en composantes principales (ACP) 172
2.2.2.1. Analyse en composantes principales sur les mesures à T1 172
2.2.2.2. Analyse en composantes principales sur les mesures à T2 174
2.2.2.3. Analyse en composantes principales sur les mesures à T3 174
2.3. Synthèse 175
2.4. Les relations prédictives entre les compétences langagières et cognitives et les épreuves de lecture 177
2.4.1. Prédicteurs de l'apprentissage de la lecture en fin de 1ère année primaire (T2) 177
2.4.1.1. Analyse des corrélations entre les épreuves à T1 et le niveau d'identification de mots à T2 . 177
2.4.1.2. Analyse des corrélations entre les épreuves à T2 et le niveau d'identification de mots à T2 . 178

2.4.1.3. Analyse de régressions pas à pas sur la variable « identification de mots écrits à T2 » 179
2.4.2. Prédicteurs de l'apprentissage de la lecture en fin de 2ème année primaire (T3)..... 179
2.4.2.1. Analyse des corrélations..... 180
2.4.2.2. Analyse de régressions multiples 183
2.4.3. Synthèse sur la contribution des différentes variables à l'explication des performances en lecture 185

3. DISCUSSION 187

CHAPITRE 3 : MODELE D'EQUATIONS STRUCTURELLES DES PREDICTEURS DE L'APPRENTISSAGE DE LA LECTURE EN ARABE 198

1. SPECIFICATION DU MODELE THEORIQUE 199

2. ESTIMATION DU MODELE DE MESURE..... 202

2.1. Estimation de la validité..... 203
2.2. Evaluation de la variance expliquée 204
2.3. Evaluation de la qualité d'ajustement 206

3. ESTIMATION DU MODELE STRUCTUREL 207

3.1. Vérification de la significativité des paramètres du modèle 209
3.2. Evaluation de la qualité d'ajustement 211

4. DISCUSSION 211

CHAPITRE 4 : EVALUATION DES ENSEIGNANTS ET PREDICTION DE LA REUSSITE ULTERIEURE EN LECTURE 215

1. METHODE 215

1.1. Sujets 215
1.2. Matériel 215
1.3. Procédure 216

2. RESULTATS 217

2.1. Caractéristiques du questionnaire renseigné par les enseignants 217
2.1.1. Corrélations inter-items..... 217
2.1.2. Analyse de la fiabilité 218
2.1.3. Analyse factorielle confirmatoire (AFC) 218
2.1.3.1. Analyse de la validité 219
2.1.3.2. Evaluation de la qualité d'ajustement..... 222

2.2. Relation prédictive entre les évaluations des enseignants en CP et les performances en lecture en 1ère année 222
2.2.1. Analyse des corrélations 223
2.2.2. Analyse de régressions multiples..... 224

2.3. Relation prédictive entre les évaluations des enseignants en CP et les compétences en lecture en 2ème année 225
2.3.1. Analyse des corrélations 225
2.3.2. Analyses de régressions multiples 226
2.3.2.1. Analyse de régressions sur les performances d'identification de mots à T3 227

2.3.2.2. Analyse de régressions sur les performances de lecture-compréhension à T3 228

2.4. Les liens prédictifs entre les évaluations des enseignants et le niveau ultérieur en lecture : proposition d'un modèle structurel 228

2.4.1. Evaluation de la qualité d'ajustement 230

2.4.2. Vérification de l'existence de relations de causalité 230

3. DISCUSSION 231

CHAPITRE 5 : VALEUR PREDICTIVE INDIVIDUELLE DES COMPETENCES LANGAGIERES ET COGNITIVES PRECOCES ET DES EVALUATIONS DES ENSEIGNANTS 237

1. CRITERES DE SELECTION 237

1.1. Sélection des enfants à partir des prédicteurs de CP et de 1ère année 238

1.2. Sélection des enfants à partir des épreuves de lecture à T1 et T2 239

2. VALEUR PREDICTIVE INDIVIDUELLE ENTRE LA CLASSE PREPARATOIRE ET LA 1ERE ANNEE 240

3. VALEUR PREDICTIVE INDIVIDUELLE ENTRE LA CLASSE PREPARATOIRE ET LA 2EME ANNEE 242

3.1. Prédiction individuelle à partir de l'épreuve de lecture IME2 242

3.2. Prédiction individuelle à partir de l'épreuve de lecture Timé2 243

3.3. Prédiction individuelle à partir de l'épreuve de lecture-compréhension 243

4. PREDICTIONS DIFFERENTIELLES 244

4.1. Comparaisons effectuées en classe préparatoire 245

4.2. Comparaisons effectuées en 1ère année 248

5. DISCUSSION 249

CONCLUSION GENERALE 253

BIBLIOGRAPHIE 271

LISTE DES TABLEAUX 296

LISTE DES FIGURES 299

ANNEXES 300

ANNEXE I : LES EPREUVES A LA FIN DE LA CLASSE PREPARATOIRE (T1) .. 301

ANNEXE II : QUESTIONNAIRE D'EVALUATION DES COMPETENCES ET COMPORTEMENTS SCOLAIRES DESTINE AUX ENSEIGNANTS 314

ANNEXE III : LES EPREUVES DE LECTURE EN 2EME ANNEE PRIMAIRE (T2) .. 320

ANNEXE IV : DEMARCHE METHODOLOGIQUE POUR LES MODELES D'EQUATIONS STRUCTURELLES 328

Avant-propos

La maîtrise de la lecture joue un rôle crucial dans le parcours scolaire et détermine en grande partie le développement professionnel, social et culturel de chaque individu. Des difficultés d'apprentissage de la lecture au cours des premières années de scolarité excluent effectivement l'enfant de la plupart des autres activités scolaires. Devant de tels enjeux, il paraît primordial de comprendre les processus fondamentaux qui sous-tendent l'apprentissage de la lecture et les raisons pour lesquelles certains enfants n'arrivent pas à bien maîtriser cette compétence. Il est en effet essentiel d'identifier les habiletés précoces qui prédisent les compétences ultérieures en lecture, ce qui est actuellement possible avec, dans des conditions précises, une fiabilité très élevée.

Une des contributions les plus importantes de ces dernières années à la compréhension des mécanismes de lecture a été la mise en évidence d'un lien étroit entre l'apprentissage du langage écrit et celui du langage oral, plus particulièrement de la capacité à réfléchir sur la structure formelle du langage (Gombert, 1990 ; 1992a ; 1992b). La capacité de segmenter les mots en phonèmes est nécessaire pour que l'enfant soit en mesure de découvrir les correspondances systématiques entre les éléments du langage parlé et écrit. La connaissance de ces correspondances est à son tour nécessaire pour reconnaître les mots rencontrés pour la première fois et pour parvenir à l'acquisition de la vitesse et de l'automaticité de l'identification des mots. Les déficits relevés dans ce domaine avant l'acquisition de la langue écrite doivent ainsi prédire la réussite ou l'échec à cet apprentissage.

Un grand nombre d'études dans le domaine de l'apprentissage de la lecture tentent de définir les facteurs qui favorisent son développement. Il est aujourd'hui largement reconnu que la conscience phonologique joue un rôle déterminant dans l'apprentissage de la lecture et apparaît comme un bon prédicteur de la réussite ultérieure en lecture (Demont & Gombert, 1996 ; Ecalle & Magnan, 2007b ; Lecoq, 1991 ; Muter, Hulme & Snowling, 2004 ; Schatschneider, Carlson & Foorman, 2004). Cependant, des travaux de recherche récents ont mis en évidence le rôle déterminant d'autres compétences comme la connaissance des lettres et les capacités de dénomination rapide (Castel, Pech-Georgel, George & Ziegler, 2008 ; Foulin, 2007 ; Schatschneider & al., 2004 ; Kirby, Parilla, Pfeiffer, 2003). En comparaison, le poids des habiletés non verbales, tout comme celui des facteurs socioculturels des enfants sont moins fortement reliés à leur niveau de lecture ultérieur que ne le sont leurs habiletés

linguistiques (Scarborough, 1989). La majorité des données disponibles concerne la langue anglaise et il paraît difficile de dire si elles sont généralisables à l'apprentissage d'autres langues alphabétiques. Les études sur les langues sémitiques sont très rares. C'est le cas de la langue arabe qui présente des propriétés linguistiques particulières.

A l'instar de de ce qui avait été montré pour la conscience phonologique, des recherches récentes montrent que les connaissances morphologiques pourraient aussi intervenir dans l'apprentissage de la lecture (Marec-Breton, 2009 ; 2003 ; Marec-Breton, Gombert & Colé, 2005). Depuis quelques années, les travaux réalisés sur l'apprentissage de la lecture en langue arabe ont mis en évidence la nécessité d'une analyse morphologique préalable au recodage grapho-phonologique (Ammar, 1997 ; 2003). La structure morphologique de l'arabe, fondée sur la racine consonantique, laisse supposer que les mots sont plus facilement décomposables en racine consonantique et morphèmes vocaliques. Les caractéristiques orthographiques de l'arabe font que les connaissances morphologiques sont nécessaires lors de la reconnaissance des mots écrits et qu'elles constituent un prédicteur puissant du niveau ultérieur en lecture (Abu Rabia, 2007 ; Boukadida, 2008).

L'évaluation des comportements et compétences scolaires par les enseignants constituent également un bon prédicteur des performances et des diffficultés scolaires ultérieures (Florin, 1991 ; Florin, Guimard & Nocus, 2002 ; Guimard & Florin, 2001 ; Guimard & Florin, 2007). L'évaluation des enseignants en classe préparatoire pourraient ainsi être utilisée de manière complémentaire aux évaluations langagières et cognitives dans une perspective de prévention des difficultés d'apprentissage (Guimard & Florin, 2008 ; Quay & Steele, 1998).

C'est dans ce contexte théorique général que s'inscrit notre travail de recherche qui visera à vérifier, en langue arabe, les résultats des études prédictives réalisées sur des langues à écriture latine. Nous espérons ainsi dégager les meilleurs prédicteurs des performances ultérieures en lecture en tenant compte des spécificités de la langue arabe et du contexte d'apprentissage dans notre pays caractérisé par la situation de diglossie.

L'objectif principal de cette étude longitudinale est d'étudier la valeur prédictive globale et individuelle des compétences langagières et cognitives précoces et des évaluations des enseignants sur les performances ultérieures en lecture. Ce travail tentera également de proposer un modèle d'équations structurelles des prédicteurs de l'apprentissage de la lecture en langue arabe. Les questions suivantes se posent donc :

Quelles habiletés précoces prédisent les performances ultérieures en lecture en langue arabe ?

Quelles sont les meilleurs prédicteurs de cet apprentissage ?

L'évaluation des comportements et compétences scolaires par les enseignants en classe préparatoire présente-t-elle un poids prédictif satisfaisant des performances ultérieures en lecture des élèves ?

Les différentes mesures utilisées dans cette recherche présentent-elles une bonne valeur prédictive individuelle ? Dans quelle mesure contribuent-elles au dépistage des difficultés de l'apprentissage de la lecture ?

Nous tentons de répondre à ces questions en nous appuyant sur l'étude d'un groupe d'enfants tunisiens suivis depuis les toutes premières étapes de l'apprentissage (classe préparatoire)[1] jusqu'à la 2ème année primaire.

Ce travail s'articule en deux grandes parties, une théorique et l'autre expérimentale. Dans la première partie, constituée de quatre chapitres, nous présenterons le cadre théorique général de la recherche.

Dans le premier chapitre, nous proposerons une synthèse de la littérature se rapportant à l'apprentissage de la lecture (identification de mots et compréhension). Nous insisterons plus particulièrement sur le développement des processus d'identification des mots écrits. Le rôle central joué par ces processus dans l'activité lexique justifie la présentation de plusieurs modèles sur cet aspect. Compte tenu des fortes spécificités de l'orthographe dans les langues sémitiques, la dernière partie sera consacrée aux études sur l'arabe et l'hébreu.

Le second chapitre propose une description des caractéristiques de la langue arabe et du fonctionnement de l'apprenti lecteur arabophone. Les travaux sur l'acquisition de la lecture en arabe ont permis d'établir certaines similitudes entre l'évolution des stratégies de l'apprenti lecteur arabophone et l'apprenti lecteur sur des langues à écriture latine. Mais, ils ont également mis en évidence des particularités, notamment au niveau des processus d'assemblage phonologique et du rôle des connaissances morphologiques dans cet apprentissage.

[1] La classe préparatoire en tunisie correspond à la grande section de maternelle (GSM) en France.

Le chapitre 3 a pour objet de présenter les principaux déterminants de la réussite en lecture en distinguant les habiletés spécifiques au traitement de l'information écrite (connaissances phonologiques, connaissances morphologiques, connaissance des lettres, etc.) et des habiletés plus générales (mémoire, niveau intellectuel, capacités visuo-attentionnelles, etc.). Nous insisterons notamment sur le rôle central de la conscience phonologique et de la conscience morphologique dans l'acquisition de la lecture. Enfin, l'influence des méthodes d'enseignement et des caractéristiques linguistiques, ainsi que le poids des facteurs sociaux sera évoquée.

Le chapitre 4 aborde la question des prédicteurs sur lesquels s'appuie l'apprentissage de la lecture. Nous présenterons une synthèse des études, réalisées sur différentes langues, qui ont essayé de déterminer les prédicteurs de réussite ou d'échec de l'apprentissage de la langue écrite. Nous distinguerons les études longitudinales prédictives, les études comparatives interlangues, les études portant sur la comparaison entre bons et faibles lecteurs et enfin les études avec entraînement métalinguistique. Nous aborderons également les travaux de recherche qui ont essayé de montrer l'importance des évaluations des enseignants dans la prédiction de la réussite scolaire ultérieure et ceux qui ont essayé de mesurer la valeur prédictive individuelle d'épreuves standardisées dans une perspective de prévention des difficultés d'apprentissage. A la fin de ce chapitre, nous metterons l'accent sur les apports et les limites des études prédictives.

La seconde partie de cette recherche, constituée de cinq chapitres, sera consacrée à la présentation et l'analyse des résultats en regard des hypothèses postulées.

Dans le chapitre 1, nous présenterons notre problématique, les objectifs et les hypothèses de la recherche. Nous exposerons également notre démarche statistique et la méthodologie adoptées pour vérifier nos hypothèses.

Le chapitre 2 explore les relations prédictives entre les compétences langagières et cognitives précoces et les performances ultérieures en lecture. La réalisation d'analyses de corrélation et de régressions multiples permettra de déterminer d'une part, la nature des relations entre les différentes mesures réalisées en classe préparatoire et les différentes épreuves de lecture et d'autre part, la valeur prédictive de ces mesures sur l'efficience ultérieure en lecture.

Dans le chapitre 3, nous proposerons un modèle structurel des prédicteurs de l'apprentissage de la lecture en langue arabe. Le recours à la méthode d'équations structurelles nous permettra

de tester ce modèle formé par les compétences évaluées en classe préparatoire considérées comme des variables explicatives et l'efficience en lecture en 2ème année, considérée quant à elle comme la variable à expliquer ou variable dépendante.

Le chapitre 4 sera consacré au rôle des enseignants de la classe préparatoire dans la prédiction du niveau de lecture des enfants deux ans plus tard. Nous avons adopté la même démarche qu'en chapitre 2 afin d'étudier la valeur prédictive globale du questionnaire d'évaluations des comportements et compétences scolaires renseigné par les enseignants sur les performances ultérieures en lecture. Nous nous sommes attachés également à étudier les qualités psychométriques de ce questionnaire.

Dans le chapitre 5, nous aborderons la question des prédictions individuelles et différentielles. Les estimations de la sensibilité, de la spécificité, de la valeur prédictive positive et négative permettront d'évaluer la valeur prédictive individuelle des mesures langagières et cognitives et du questionnaire renseigné par les enseignants.

En conclusion de notre travail, nous propserons une synthèse des différents points abordés tout en mettant l'accent sur les contributions et les limites de cette recherche.

Première partie :
Cadre théorique de la recherche

Chapitre 1 : L'apprentissage de la lecture

L'apprentissage de la lecture constitue un des apprentissages les plus importants des premières années de scolarité. L'échec dans cet apprentissage peut avoir des causes différentes et est très souvent avancé comme explication de la non-réussite scolaire. La perspective cognitive est à l'origine de points de vue et d'approches qui ont permis des avancées considérables dans la recherche sur la lecture. Celle-ci est conçue comme une habileté complexe impliquant plusieurs processus et ressources perceptives, cognitives et linguistiques diverses. Depuis plus d'une trentaine d'années, les chercheurs se sont intéressés à l'étude des procédures mises en place lors de l'apprentissage de la lecture. La façon dont les jeunes lecteurs apprennent à lire et dont les lecteurs experts traitent l'information écrite représente pour la psychologie cognitive un enjeu important et fait l'objet de recherches très fructueuses. L'étude des différentes procédures impliquées dans l'apprentissage de la lecture et leurs modélisations fournit ainsi un cadre pour la compréhension des difficultés en lecture en permettant de situer les dysfonctionnements responsables de ces difficultés.

L'objectif de ce chapitre est de décrire les procédures d'identification des mots écrits et de compréhension chez le lecteur débutant et chez le lecteur expert. Nous présenterons ensuite les principaux modèles de l'apprentissage de la lecture en distinguant les modèles développementaux, les modèles par analogie et les modèles interactifs. Enfin, la dernière partie sera consacrée aux études sur l'apprentissage de la lecture en arabe et en hébreu.

1. L'identification des mots écrits

1.1. Définition

De nombreux auteurs considèrent la lecture comme une activité complexe de traitement des informations graphiques afin d'en extraire la signification (*cf.* Perfetti, 1989 ; Morais, 1994). Plus particulièrement, la lecture est conçue comme le produit de l'intégration de deux étapes considérées généralement comme des composantes : une composante de traitement de bas niveau, la reconnaissance des mots écrits et une composante de plus haut niveau, la compréhension. La compétence en lecture est donc définie comme la combinaison de deux composantes nécessaires dont aucune n'est à elle seule suffisante. Cependant, si les deux types de processus doivent être efficients, de nombreuses études ont montré que les difficultés

dans la reconnaissance des mots écrits constituaient la source principale des échecs en lecture (Blanc, 2009 ; Alegria & Morais, 1996 ; Gombert & Fayol, 1995 ; Morais, 1994). L'objectif principal de l'apprenti lecteur est de parvenir à comprendre ce qu'il lit de la même façon qu'il comprend ce qu'il entend. Dans ce but, celui-ci doit, dans un premier temps, automatiser l'identification des mots écrits et donc développer des procédures de lecture rapides, précises et automatiques. De cette façon, une partie importante de la capacité de traitement pourra être orientée vers la compréhension de phrases ou de textes. L'habileté à décoder et la capacité de comprendre prennent une part plus ou moins importante en fonction du niveau atteint par le lecteur. Ainsi, d'un point de vue développemental, plus l'élève progresse dans sa scolarité, moins il fait appel au décodage. Développer des mécanismes efficaces de reconnaissance de mots écrits est un aspect essentiel de l'acquisition de la lecture (Gombert & Fayol, 1995). En effet, dans une langue alphabétique, comme le français et l'arabe, l'automatisation de l'habileté à décoder revêt une grande importance pour le lecteur débutant.

Morais (1994) précise que « *ce qui existe de spécifique à l'activité de lecture est la capacité de reconnaissance des mots écrits, c'est-à-dire la capacité d'identifier chaque mot en tant que forme orthographique ayant une signification et de lui attribuer une prononciation* » (p. 116). Il parait, en effet, inconcevable d'accéder à la signification du texte sans pour autant identifier au moins une partie des mots qu'il comporte. Stanovich (1989) confirme que des problèmes au niveau du processus d'identification des mots sont un facteur critique dans la plupart des cas de dyslexie. L'efficacité de ce processus rend compte d'une grande partie de la variance de la performance en lecture. En effet, la précision et la rapidité de l'identification des mots conditionnent l'activité de lecture. C'est donc essentiellement à l'identification des mots que s'attache notre recherche.

Une idée commune à tous les travaux consacrés à l'identification des mots est de supposer l'existence d'un lexique mental construit au cours des premières années d'acquisition du langage et dans lequel sont stockés, de manière stable, l'ensemble des mots connus et utilisés (Seymour, 1986). Ce lexique mental se présente comme un ensemble de représentations correspondant aux unités signifiantes de la langue (Caron, 1989). Identifier un mot écrit consiste donc à mettre en rapport la représentation écrite du mot avec sa représentation interne correspondante. Selon Zagar (1992), le lexique mental « *se présente comme un dictionnaire interne composé de tous les mots connus auxquels sont associées les informations orthographiques, phonologiques, syntaxiques et sémantiques qui lui sont propres* » (p. 49). Dans cette optique, identifier un mot consiste à chercher son entrée lexicale dans un

dictionnaire de langue ; une fois trouvée, on sait comment ce mot se prononce et ce qu'il signifie. L'identification ou la reconnaissance de mots[2] renvoie à la capacité d'établir une association entre une forme graphique extraite du stimulus et une représentation particulière récupérée en mémoire à partir du lexique mental. Selon Forster (1976, cité par Zagar, 1992), l'identification des mots s'effectue de la façon suivante :

- A partir du mot lu est élaborée une description codée du stimulus, c'est l'étape ou s'effectue l'extraction des signes et la construction du code d'entrée au lexique.
- Ensuite, une recherche active est effectuée dans le fichier d'accès orthographique pour recouvrer l'information lexicale, c'est l'étape d'accès au lexique proprement dite.
- Enfin, le processus se termine par une compréhension détaillée des correspondances entre les informations concernant les mots et les propriétés du stimulus, c'est l'étape de vérification post-accès. Cette dernière étape concerne le recouvrement de la signification du mot.

1.2. L'identification des mots écrits chez l'apprenti lecteur

Dans les systèmes d'écriture alphabétique, l'existence d'une relation entre la représentation graphique et les caractéristiques de surface de la langue parlée a conduit à s'interroger sur le rôle possible de l'information phonologique dans l'identification de mots chez l'apprenti lecteur.

1.2.1. La médiation phonologique

L'objectif principal de l'apprenti lecteur est de développer des procédures d'identification de mots rapides, précises et automatiques. De cette façon, une partie importante de la capacité de traitement pourra être orientée vers la compréhension de phrases ou de textes (Colé & Fayol, 2000). Ces procédures d'identification de mots automatiques consistent principalement en une série de traitements impliquant l'analyse visuelle de la structure formelle des mots et les traitements phonologiques et orthographiques subséquents. Parmi ces processus, il convient

[2] Certains auteurs, comme Sprenger-Charolles & Colé (2003), pensent qu'il est important d'opérer une distinction entre identification des mots et reconnaissance des mots. L'identification des mots écrits renvoie à l'ensemble des procédures de décodage du mot-stimulus qui permet d'en activer la représentation en mémoire alors que leur reconnaissance consiste à lire les contenus de la représentation activée en mémoire pour accéder au sens de ce qui est lu. De ce point de vue, l'identification des mots peut être considérée comme l'étape préalable à leur reconnaissance. Dans cette recherche, nous considérons les deux termes comme équivalents.

de souligner l'importance de la médiation phonologique qui joue un rôle central dans les processus d'identification de mots écrits.

Une caractéristique essentielle des écritures alphabétiques, que l'on a parfois appelé le principe alphabétique, est que les signes écrits élémentaires - les lettres - correspondent de façon stable à des sons élémentaires de la langue parlée. L'idée de médiation phonologique s'inscrit dans cette perspective. Pour apprendre à lire, l'enfant doit développer des connaissances sur les liens entre les segments du langage parlé et les segments de l'écrit (Alegria & Morais, 1994 ; Ecalle & Magnan, 2002 ; Goswami & Bryant, 1990). La notion de médiation phonologique renvoie à l'idée selon laquelle, au cours de la reconnaissance visuelle d'un mot, le lecteur construit une représentation mentale de la façon dont le mot s'entend et se prononce. La reconnaissance des mots repose sur l'exploitation des règles de correspondances entre les unités visuelles discrètes (les graphèmes) et des unités phonémiques abstraites (les phonèmes). L'apprenti lecteur doit donc comprendre le principe alphabétique, ce qui lui permet alors d'utiliser une procédure de lecture phonologique qui consiste à traduire la séquence de lettres du mot lu en une séquence de sons qui correspondent à ces lettres. Les arguments en faveur d'une telle médiation sont nombreux. D'abord, l'apprentissage de la lecture exploite une partie des processus déjà établis pour le langage parlé (Zagar, 1992 ; Peerman, 1991). Ensuite, certains auteurs (par exemple Casalis & Lecoq, 1992 ; Casalis, 1995 ; Sprenger-Charolles & Colé, 2003) ont pu observer des cas de dyslexie qui mettent en évidence une atteinte spécifique de la voie phonologique. Mais la raison principale motivant cette hypothèse réside dans notre habileté à prononcer des séquences de lettres ne correspondant pas à un mot (pseudo-mot) et par conséquent non représentées dans le lexique mental.

La médiation phonologique joue donc un rôle extrêmement important dans l'acquisition de la lecture. Dans un premier temps, une représentation phonologique du mot serait utilisée pour accéder aux représentations lexicales. Ensuite, l'appariement de la représentation assemblée du mot avec sa représentation phonologique stockée dans le lexique mental conduirait à l'identification du mot. La connaissance des correspondances entre graphèmes et phonèmes ne suffit cependant pas à une maîtrise des mots écrits. Certains mots comportent, en effet, des irrégularités qui les rendent difficiles à décoder par stricte application des correspondances grapho-phonologiques (par exemple : femme, monsieur…). Par ailleurs, il existe des mots homophones qu'une procédure purement phonologique ne permettrait pas de les distinguer de point de vue de la signification (exemple : chant et champs ; conte et compte…). La

reconnaissance de ces mots est effectuée directement par l'analyse visuelle. Le recours systématique à la conversion phonologique constitue un puissant mécanisme d'auto-apprentissage en permettant au lecteur débutant d'identifier des mots rencontrés pour la première fois ou, plus généralement, des mots pour lesquels une représentation orthographique n'est pas disponible (Sprenger-Charolles & Colé, 2003).

1.2.2. Rôle du contexte dans l'identification des mots

Les recherches actuelles convergent pour montrer qu'un des enjeux premiers de l'apprentissage de la lecture est l'installation de procédures de reconnaissance des mots écrits qui devront, à terme, être automatiques. Il a été maintes fois montré que les faibles lecteurs ont généralement des performances faibles dans les tâches de décodage qui impliquent l'application des règles de correspondance grapho-phonologiques (Alegria & Morais, 1989 ; Doctor & Coltheart, 1980 ; Sprenger-charolles, 1992 ; Sprenger-Charolles & Casalis, 1996). Mais de surcroît, ils semblent également moins habiles dans l'application de stratégies consistant à utiliser le contexte pour identifier un mot. Très tôt, l'enfant pré-lecteur est capable de reconnaître des mots écrits présentés dans leur contexte familier (par exemple sur des étiquettes ou des enseignes). Il semble, dans ce cas, que ce soit uniquement le contexte et non pas l'écrit lui-même qui détermine l'interprétation du jeune enfant (voir par exemple Masonheimer, Drum & Ehri, 1984 ; Sprenger-Charolles & Casalis, 1996). Ainsi, la plupart de ces enfants sont incapables de reconnaître un mot présenté dans le contexte habituel d'un autre mot (XEPSI pour PEPSI par exemple).

De nombreuses recherches ont comparé les stratégies utilisées par les lecteurs débutants, normaux ou retardés, en ce qui concerne l'utilisation du contexte pour l'identification d'un mot. Les résultats obtenus par Sprenger-Charolles & Khomsi (1989) indiquent que l'utilisation du contexte dans l'identification d'un mot est surtout observée chez les lecteurs les moins performants. Les auteurs suscitent les effets du contexte par une image accompagnant un mot dont le sujet doit simplement dire s'il est le bon mot. Certains mots correspondent à l'image et leur forme graphique est correcte, d'autres présentent une perturbation graphique (exemple : *autopobile* pour *automobile*) ; enfin, certains mots présentent une perturbation sémantico-contextuelle (*limace* sous l'image d'un *escargot*). Les bons lecteurs détectent toutes les perturbations graphiques et la majorité des perturbations sémantico-textuelles alors que les faibles lecteurs détectent 81% des perturbations sémantico-

contextuelles (détections qui impliquent l'usage du contexte) et seulement 37% des perturbations graphiques.

L'expérience de Perfetti, Goldman & Hogoboam (1979) menée avec des enfants de dix ans, bons et faibles lecteurs, confortent ces résultats. Ces enfants devaient lire des mots qui sont ou non intégrés dans une histoire. Les résultats montrent que cette intégration facilite la lecture des mots pour les deux groupes de sujets et que cette facilitation est beaucoup plus importante dans le groupe des faibles lecteurs. Les auteurs observent également que les faibles lecteurs manifestent des effets de contexte quelque soit la fréquence des mots écrits alors que pour les bons lecteurs, un tel effet n'est observé que pour les mots peu fréquents. Ainsi, non seulement les stratégies de devinette à partir du contexte sont généralement des stratégies de débutants ou de faibles lecteurs mais il semble que les faibles lecteurs sont moins équipés que les autres pour les adopter. Le fait que les bons lecteurs sont moins sensibles au contexte que les faibles lecteurs dans la reconnaissance des mots montre que le développement des habiletés de lecture va dans le sens d'un traitement de plus en plus indépendant, automatisé, précis et rapide de chaque mot. Selon Perfetti (1989), le contexte confirme la signification extraite des mots mais n'influence que marginalement l'identification elle-même. Ce point est fondamental pour rendre compte des processus d'identification rapides et indépendants du contexte décrits chez le lecteur expert et des processus d'identification lents et dépendants du contexte décrits chez le lecteur moins accompli.

Par ailleurs, il a été montré que la plupart des mots ne sont pas prédictibles par le contexte et que les mots les moins prédictibles sont justement les plus informatifs et les plus difficiles à identifier (Gough & Juel, 1989). Le contexte permet essentiellement d'aider la reconnaissance des mots courts et familiers, précisément ceux dont la reconnaissance pose le moins de problèmes. Nicholson (1991) montre que l'utilisation du contexte pour reconnaître les mots écrits est dépendante du niveau d'expertise en lecture. Dans ce but, il compare le pourcentage de gain en lecture de mots lus en contexte phrasique et de mots lus de façon isolée chez des enfants âgés de 6, 7 et 8 ans et de différents niveaux de lecture (mauvais, bons et normolecteurs). Les résultats indiquent des effets bénéfiques de la lecture en contexte phrasique uniquement pour les mauvais lecteurs de chaque âge et pour les normolecteurs de 6 et 7 ans. Ainsi, chez le lecteur habile, l'effet facilitateur du contexte phrasique est très faible sinon inexistant. Dans le même sens, Stanovich (1979) demande à des enfants et des adultes de lire un mot qui leur est présenté sur un écran à la suite d'une phrase inachevée. Certains de ces mots sont congruents avec la phrase et l'effet facilitateur du contexte peut donc intervenir

dans leur lecture, d'autres sont non congruents. Il constate que les sujets adultes lisent aussi vite les deux catégories de mots alors que chez les sujets de 9 à 11ans, le temps de latence précédant la lecture est plus long pour les mots non congruents. Ce sont donc les lecteurs les moins expérimentés qui se réfèrent le plus au contexte.

Contrairement aux données recueillies sur l'anglais et le français, les résultats disponibles en langue arabe montrent que les élèves faibles lecteurs utilisent le contexte à un moindre degré que les bons lecteurs (Abu-Rabia, 2001 ; 1997 ; 1996 ; Abu-rabia & Siegel, 1995 ; Ammar, 1997). Ces différences sont dues à la structure linguistique propre à chaque langue. Nous présenterons ces études dans le chapitre suivant.

1.3. L'identification de mots écrits en lecture experte

Apprendre à lire consiste à développer une voie d'accès aux connaissances linguistiques à partir de la représentation écrite de la langue (Ecalle & Magnan, 2002). Un aspect essentiel de cet apprentissage consiste à développer des mécanismes efficaces d'identification de mots. Avant l'apprentissage de la lecture, les enfants ne possèdent pas de lexique orthographique et ont donc recours principalement à une procédure phonologique pour lire les mots. L'utilisation de cette procédure témoigne de la compréhension du principe fondamental d'une écriture alphabétique dont le but est de représenter les unités phonémiques de la langue orale. Le rôle central de la médiation phonologique au début de l'acquisition de la lecture a été mise en évidence durant ces trente dernières années dans de nombreux travaux et fait aujourd'hui l'objet d'un consensus (Alegria & Morais, 1994 ; Demont & Gombert, 2007 ; Gombert, 1990 ; Ferrand, 1995 ; Peerman, 1991 ; Sprenger-Charolles & Casalis, 1996). Cependant, la médiation phonologique utilisée au début de l'apprentissage de la lecture diffère de celle utilisée par le lecteur expert notamment par le fait qu'elle est en développement.

Il est classique de considérer que l'aboutissement du développement de la compétence en lecture est un fonctionnement expert compatible avec le modèle à double voie. Ce modèle décrit deux procédures indépendantes d'accès au lexique. La première est une procédure directe par adressage qui permettrait la reconnaissance des mots par appariement direct de la configuration écrite du mot avec sa représentation visuelle en mémoire. La seconde procédure est indirecte, elle consiste à dériver une représentation de nature phonologique à partir de la traduction de segments orthographiques (lettres ou groupes de lettres) en segments phonologiques (phonèmes ou groupes de phonèmes). Ainsi, pour reconnaître les mots écrits, le lecteur expert utilise soit des informations orthographiques issues du traitement perceptif du

mot à identifier pour activer directement une représentation en mémoire stockée dans le lexique orthographique, soit des unités infra-lexicales pour associer les patterns orthographiques des mots écrits à des patterns phonologiques correspondants (Ecalle & Magnan, 2002). L'analyse expérimentale des mécanismes de lecture effectivement utilisés par le lecteur habile montre que les deux procédures interviennent en parallèle : celle d'accès direct serait à l'œuvre dans l'identification de mots familiers, celle d'accès indirect conduit à l'identification des mots non familiers et des mots irréguliers. La question actuellement étudiée est plutôt de déterminer dans quelle mesure l'information phonologique extraite des séquences de lettres contribue à la reconnaissance des mots chez le lecteur expert.

1.3.1. L'assemblage phonologique chez le lecteur expert

L'intervention du code phonologique dans le processus d'identification des mots a été mise en évidence dans les expériences utilisant la tâche de décision lexicale. Dans cette tâche, le sujet est invité à juger si les séquences de lettres qui lui sont présentées sont des mots de la langue ou des pseudo-mots. Contrairement aux mots, les pseudo-mots ne peuvent être prononcés qu'en recourant à la procédure d'assemblage. Les travaux portant sur les effets d'homophonie avec des mots et des pseudo-mots montrent que les lecteurs produisent plus d'erreurs et manifestent des temps de rejet plus lents pour les pseudo-homophones (exemple *brane* pour *brain*) que pour les pseudo-mots contrôle (Coltheart, Davelaar, Jonasson & Besner, 1977 ; Stone & Van Orden, 1993, cités par Sprenger-Charolles & Colé, 2003).

Dans le même sens, Van Orden (1987) a décrit un effet d'homophonie des pseudo-mots dans une tâche de catégorisation sémantique. Dans cette tâche, le sujet est invité à juger si la séquence de lettres présentée est membre ou non de la même catégorie sémantique prédésignée. Les résultats ont montré que, tout comme pour les pseudo-mots, il est plus difficile de juger qu'un mot n'appartient pas à une catégorie sémantique lorsque ce mot est homophone d'un mot membre de cette catégorie. Ainsi, les sujets adultes éprouvent plus de difficultés à rejeter *Rows*, homophone incorrect de *Rose*, que l'item contrôle *Robs* comme exemplaire de la catégorie des fleurs. D'après l'auteur, ces données témoignent de l'utilisation d'une représentation phonologique assemblée dans l'identification de mots. Le code phonologique de *Rows* activerait l'entrée lexicale du mot *Rose* d'où une ambiguïté source d'erreurs. La seconde observation effectuée avec des mots en décision lexicale concerne la régularité grapho-phonologique qui traduit le fait que certains mots possèdent une prononciation irrégulière imprévisible sur la base de leur orthographe (par exemple femme,

monsieur). Ainsi, les processus d'assemblage phonologique donneraient lieu à une représentation phonologique correcte pour les mots réguliers mais pas pour les mots irréguliers. Les données indiquent que l'effet de régularité ne survient que pour les mots d'usage peu fréquent (Waters, Seidenberg & Bruck, 1984). Chez le lecteur expert, la représentation phonologique des mots serait plus rapidement obtenue par le recours à l'information stockée lexicalement que par l'intermédiaire des processus d'assemblage. Il semble, toutefois, que lorsque l'identification des mots à partir de l'information orthographique nécessite plus de temps (mots peu fréquents), la représentation phonologique assemblée pourrait devenir disponible à temps pour intervenir dans le processus d'identification de ces mots.

L'utilisation d'un code phonologique en reconnaissance de mots a également été mise en évidence par la technique d'amorçage (présentation d'un mot ou d'un pseudo-mot amorce qui précède celle d'un mot cible). L'intérêt de l'amorçage réside principalement dans la possibilité de manipuler la nature de la relation entre l'amorce et la cible afin d'étudier les différents niveaux de traitements impliqués en lecture. Ainsi, amorce et cible peuvent entretenir des relations visuelles, orthographiques, phonologiques et sémantiques. La question est de savoir si une amorce à un effet facilitateur ou inhibiteur sur le traitement de la cible. Dans des expériences réalisées en français par Ferrand & Grainger (1993, 1994), les sujets doivent effectuer une tâche de décision lexicale ou de dénomination sur un mot cible précédé par la présentation d'un mot amorce pseudo-homophone (exemple : piez/PIED) ou par celle d'un pseudo-mot amorce orthographiquement relié (exemple : pien/PIED) ou encore par celle d'un pseudo-mot non relié (exemple : bate/PIED). Amorce et cible sont présentées dans des typographies différentes pour éviter que la tâche ne soit effectuée sur la base d'indices visuels uniquement. Les auteurs constatent que l'identification du mot cible est meilleure lorsque celui-ci est précédé d'une amorce phonologique ou orthographique comparativement à une amorce contrôle.

Les résultats de ces expériences ont montré, d'une part, que pendant la reconnaissance des mots écrits, l'activation des codes orthographiques et phonologiques est extrêmement rapide et, d'autre part, que l'activation des codes orthographiques est plus rapide que celle des codes phonologiques. Les effets de facilitation phonologique sont observés quelle que soit la fréquence des mots écrits, ce qui suggère que le lecteur expert active systématiquement les codes phonologiques des mots pendant leur reconnaissance.

1.3.2. L'importance de l'automatisation des procédures d'identification de mots

La procédure phonologique, systématiquement utilisée par les enfants au début de l'apprentissage de la lecture, s'apparente à la voie indirecte décrite dans les modèles à double voie. Cependant, la procédure d'assemblage phonologique du lecteur habile ne doit pas être confondue avec le déchiffrage lent et laborieux du lecteur débutant ou faible lecteur (Alegria & Morais, 1994 ; Demont & Gombert, 2004). Dans le premier cas, l'activité en jeu est entièrement automatique et requiert peu ou pas de ressources attentionnelles alors que, dans le second, elle exige une attention soutenue de la part du lecteur, donc que celui-ci mobilise vers cette activité une partie plus ou moins importante de ses capacités cognitives. Cette allocation de l'attention à l'activité de l'identification des mots s'opère au détriment des processus de compréhension. Dans un tel cas, l'élève n'a pas la possibilité de se préoccuper du sens du texte.

Lorsque l'enfant est confronté à l'écrit, il doit, d'une part, stocker de nouvelles informations dans son lexique mental et, d'autre part, créer de nouvelles voies d'accès (Sprenger-Charolles, 1992). L'identification doit être, non seulement correctement effectuée, mais également automatisée. Un bon niveau d'automatisation de ce processus est en effet indispensable pour permettre au lecteur de consacrer le maximum de ressources cognitives aux traitements nécessaires à la compréhension. Chez le lecteur habile, les activités d'identification et de compréhension se dérouleraient simultanément et en interaction. Pour le jeune lecteur et/ou le faible lecteur, en revanche, les procédures d'accès au lexique n'étant pas encore automatisées, les deux types d'activité ne pourront pas en conséquence se dérouler simultanément. De fait, apprendre à lire consiste à acquérir des « réflexes » de lecture qui requièrent un haut niveau d'automaticité (Sprenger-Charolles & Casalis, 1996).

Dès la fin de la première année, l'identification des mots écrits repose de plus en plus sur des processus automatisés et ne requièrent plus d'importantes ressources attentionnelles. En conséquence, l'automatisation progressive des processus d'identification des mots permet de diminuer la charge attentionnelle nécessaire au décodage et d'allouer le maximum de ressources cognitives aux aspects sémantiques et syntaxiques de la phrase et/ou du texte. La compréhension de phrases ou de textes s'effectue ainsi aisément et sans effort cognitif apparent chez le lecteur expert puisqu'elle s'appuie sur une identification des mots écrits qui se déroule de façon quasi-réflexe (Sprenger-Charolles & Colé, 2003). Guttentag & Haith

(1978) ont présenté à des adultes des images représentant des animaux, des meubles et des moyens de transport à l'intérieur desquels se trouve un mot écrit. Ce mot appartient soit à la même catégorie sémantique que le dessin, soit à une catégorie différente. La tâche des sujets consiste à dénommer, le plus rapidement possible, le nom du dessin en ignorant le mot écrit à l'intérieur. L'effet d'interférence obtenu entre les deux conditions catégorielles (effet stroop) suggère que les procédures d'identification de mots écrits se déroulent en dehors de tout contrôle attentionnel. L'automatisation des processus à l'œuvre dans la reconnaissance des mots libère les ressources cognitives pour les activités de plus haut niveau. Ainsi, le lecteur expert peut consacrer la quasi-totalité de son effort de réflexion aux informations apportées par le texte.[3] La lecture experte correspond à l'automatisation des procédures phonologique et orthographique permettant de reconnaître les mots avec précision et rapidité. La procédure orthographique est généralement utilisée pour identifier de manière directe les mots familiers et les mots irréguliers ; la procédure phonologique serait, quant à elle, utilisée pour identifier les mots inconnus ou les pseudo-mots. La médiation phonologique est considérée comme un mécanisme de secours, relativement lent, et servant essentiellement à décrypter des mots dont la forme écrite n'est pas familière ou plus généralement des mots pour lesquels une représentation orthographique stockée n'est pas disponible. L'analyse des processus d'identification des mots permet de caractériser le lecteur expert de deux manières. D'une part, il a développé une procédure d'assemblage efficace et entièrement automatisée. D'autre part, il possède un large corpus de représentations des mots permettant leur identification directe.

Selon Gombert (2003), deux types de processus peuvent expliquer l'installation des automatismes de lecture caractéristiques de la lecture experte : l'automatisation et l'apprentissage implicite. Dans la perspective de l'automatisation, la lecture, initialement gérée consciemment, s'automatise par répétition de l'activité de lecture. Cette répétition permet le désengagement de l'attention et l'accélération du traitement. Dans le cadre d'un apprentissage implicite, les régularités internes et contextuelles présentées par les mots perçus affectent progressivement la pondération de connexions internes au système cognitif, ce qui se traduit par une évolution des réponses évoquées automatiquement lors de la confrontation aux mots écrits. Apprendre à lire consiste ainsi à développer et automatiser les composantes spécifiques à la lecture.

[3] Selon Gombert & Fayol (1995), cette automatisation est dépendante de l'exercice de la répétition. Le seul moyen pour automatiser la lecture est de lire. C'est ce qui explique qu'avec l'âge, le fossé se creuse entre bons et mauvais lecteurs (plus on lit, mieux on lit !).

2. La compréhension en lecture

Dans le modèle de Gough & Tumner (1986), la performance en lecture (L) est présentée comme la résultante de la combinaison de deux variables, la reconnaissance des mots (R) et la compréhension orale (C) : soit L= R x C. Il est en effet impossible d'accéder à la signification d'un texte sans identifier au moins une partie des mots qu'il comporte. Si l'identification correcte des mots écrits est nécessaire, elle n'apparait pas cependant suffisante et ne conditionne pas à elle seule l'activité de lecture dont la finalité est bien la compréhension de ce qui est lu (Blanc, 2009 ; Gombert & Fayol, 1995).

Les études de corrélations entre les performances en identification de mots et en compréhension de textes révèlent l'existence de liaisons très fortes dont l'intensité diminue avec le niveau scolaire (Juel, Griffith & Gough, 1986). Le traitement rapide et correct des mots écrits constitue ainsi un facteur crucial dans la compréhension de textes (Stanovich, 1986). Cependant, à un niveau de décodage égal, certains enfants présentent plus que d'autres des déficits de compréhension qui se traduisent par leur faible capacité à répondre à des questions relatives au contenu des textes qu'ils ont lus (Blanc, 2009). D'autres dimensions interviennent, notamment celles qui ont trait aux capacités de traitement, aux mécanismes d'inférences et aux habiletés des lecteurs à contrôler, évaluer et réguler leurs propres habiletés de traitement.

2.1. Les mécanismes cognitifs de la compréhension

Outre le décodage, la compréhension en lecture suppose un contrôle métalinguistique et une autorégulation autonome (Gombert & Fayol, 1995). Elle exige une capacité à prendre en compte d'une part la syntaxe des phrases lues ; d'autre part le détail des informations explicites et implicites apportées effectivement par chaque texte et enfin le contrôle et la régulation de la compréhension (Fayol, 1992). La compréhension en lecture met en jeu des processus locaux avec traitement des indices linguistiques (lexicaux, morphologiques et syntaxiques) au cours de la lecture des mots et une activité de plus haut niveau engagée dans la production d'une situation mentale construite au cours de la lecture du texte (Lecoq, Casalis, Leuwers & Watteau, 1996). Il existe des processus orientés vers la compréhension des éléments de la phrase, d'autres vers la recherche de cohérence entre les phrases, d'autres encore ont comme fonction de construire un modèle mental du texte ou une vision d'ensemble qui permettra au lecteur d'en saisir les éléments essentiels et par la suite de faire des

hypothèses, d'intégrer le texte à ses connaissances antérieures, d'autres processus enfin servent à gérer la compréhension (Blanc, 2009).

Pour comprendre le texte, le lecteur doit effectuer une analyse syntaxique, ce qui demande au minimum que soit prise en compte la structure grammaticale qui gouverne l'organisation des mots dans la phrase et que les marques de cohésion textuelle (qui lient les propositions et les phrases entre elles) soient traitées. Lire une phrase, c'est identifier les mots et en même temps reconnaître leurs rôles grammaticaux respectifs. La conscience syntaxique facilite les processus d'intégration de texte et rend le lecteur capable de contrôler plus efficacement ses processus de compréhension (Demont, 2002). Le processus qui consiste à donner un rôle aux mots dans la phrase est nommé calcul syntaxique. Le calcul syntaxique de la signification de la phrase est ainsi conçu comme une étape supplémentaire de traitement en lecture, au-delà de l'accès au lexique.

Les études longitudinales confirment les liens entre conscience syntaxique et lecture-compréhension (Demont & Gombert, 1996 ; Nocus & Gombert, 1997). Nocus & Gombert (1997) ont montré, dans une étude longitudinale, que les performances dans les tâches de reproduction de fautes de syntaxe suivant un modèle, mesuré au CP, sont les meilleurs prédicteurs des performances dans un test de compréhension l'année suivante, et ceci après contrôle du QI, du vocabulaire et de l'empan de la mémoire de travail. Pour comprendre un texte, le lecteur doit en identifier les mots, mettre en relation les mots et les phrases pour construire une représentation mentale globale et cohérente de la situation décrite par le texte. La compréhension fine des messages nécessite d'aller au-delà de ce qui est dit dans le texte ou l'énoncé, c'est-à-dire de pouvoir faire le lien entre l'information explicitée par le texte et les connaissances qu'on a déjà dans notre mémoire à long terme : c'est le mécanisme d'inférence (Lieury & De La Haye, 2004). Il s'agit de se représenter des personnages, des lieux, des actions et des évènements qui ne sont pas tous mentionnés explicitement dans le texte. Selon Lieury & De La Haye (2004), « *Les inférences jouent un rôle primordial dans la compréhension, car elles participent activement et pour une large part à la construction du modèle de situation et permettent entre autre de réduire l'ambiguïté lexicale, de résoudre les références anaphoriques, d'identifier le contexte, de reconnaître les événements non congruents ou encore de prédire les causes et les conséquences des événements.* » (p. 75). La signification n'est pas donnée par le texte mais elle est construite par le lecteur et varie donc autant en fonction de la base de connaissances et des stratégies du lecteur qu'en fonction de l'information apportée (Gombert & Fayol, 1995). Comprendre suppose donc non seulement

d'identifier les informations explicites d'un texte mais également de rendre explicites les informations implicites qu'il recèle (De La Haye & Bonneton-Botté, 2009).

2.2. Facteurs de difficulté en compréhension

La compréhension d'un texte peut échouer ou présenter des difficultés pour différentes raisons. La première cause de difficulté en compréhension est liée à un déficit dans les processus d'identification de mots écrits. Des processus trop lents, imprécis et peu automatisés entravent l'accès à la compréhension. La compréhension en lecture dépend globalement de l'intervention de deux types d'habiletés, celles dont dépend le décodage des mots que sont les habiletés langagières de base (conscience phonologique, connaissance des lettres) et celles dont dépend la compréhension de la situation décrite (produire des inférences). Selon Cain, Oakhill & Bryant (2000, cité par Ecalle, Magnan & Ramus, 2007), les faibles compreneurs se distinguent par des difficultés au niveau de la mise en œuvre de processus de haut niveau et non pas par des difficultés au niveau des compétences phonologiques comparativement aux bons compreneurs. Une des différences essentielles entre les bons et les mauvais compreneurs est cette difficulté à produire des inférences. Les mauvais compreneurs réussissent difficilement les tâches qui nécessitent un traitement de l'information et l'élaboration d'inférences (De La Haye & Bonneton-Botté, 2009). Cette difficulté s'étend à l'extraction de la signification des mots à partir du contexte et à l'extraction des liens qui donnent cohérence à une histoire.

Cain & Oakhill (1999, cité par Blanc, 2009) ont mené une étude ayant pour objectif de déterminer le sens de la relation entre la production d'inférences et les habiletés de compréhension. Quatre récits de longueurs différentes ont été utilisés. Trois groupes étaient ainsi comparés quant à leurs productions inférentielles : bons compreneurs, mauvais compreneurs et groupe contrôle. Les résultats montrent que les mauvais compreneurs obtenaient des performances semblables à celles des deux autres groupes pour les questions qui portaient sur les informations explicitement fournies dans le texte, alors qu'ils obtenaient de moins bonnes performances que les deux autres groupes pour les questions qui nécessitaient la production d'inférences. D'après les auteurs, cette moindre propension des mauvais compreneurs à produire des inférences peut être expliquée par le fait qu'ils adoptent des objectifs de lecture qui sont différents de ceux que suivent les bons lecteurs. Il semble, en effet, que les mauvais compreneurs se focalisent plus sur la précision de lecture des mots que sur l'auto-évaluation de leur propre compréhension de la situation évoquée.

Les bons et les mauvais compreneurs se distinguent également par leur capacité d'autorégulation ou la capacité de contrôler sa propre compréhension (Goigoux, 2000). Certains mauvais compreneurs éprouvent en effet des difficultés à se rendre compte de la défaillance de leur compréhension et à mettre en œuvre les stratégies adéquates pour remédier aux difficultés éventuelles : réguler sa vitesse de lecture, effectuer des retours en arrière dans le texte, porter plus d'attention aux parties qui semblent plus ardues, etc. D'autres travaux insistent sur le lien entre compréhension en lecture et mémoire de travail. La mémoire de travail joue un rôle important dans le traitement de phrases syntaxiquement complexes et dans le processus de sélection et de hiérarchisation des informations traitées dans les phrases au cours de la lecture du texte (Ecalle, Magnan & Ramus, 2007). La compréhension en lecture nécessite de maintenir l'information en mémoire pendant le traitement simultané des autres informations. Les lecteurs avec une capacité déficitaire de mémoire de travail devraient avoir moins de ressources disponibles pour le maintien de l'information lors de la compréhension (Demont, 2002).

Enfin, le poids des connaissances lexicales joue un rôle dans les performances de compréhension. Tannenbaum, Torgesen & Wagner (2006, cités par Ecalle & al., 2007) ont étudié les liens entre les connaissances lexicales et la compréhension écrite. Les auteurs ont évalué les connaissances lexicales en distinguant trois dimensions : l'étendue du vocabulaire (taille), sa profondeur (richesse) et la fluence (vitesse d'accès au sens). Les résultats montrent que ces trois dimensions rendent comptent de 62% de la variance en compréhension en lecture. Par ailleurs, des études longitudinales récentes supportent l'idée selon laquelle les habiletés de compréhension évaluées chez les enfants d'âge préscolaire permetteraient de prédire leurs compétences en matière de compréhension à mesure qu'ils avancent en âge (Kendeou & al., 2005 ; Van den Broek & al., 2005 cités par Blanc, 2009). Ainsi, les habiletés précoces de compréhension sont jugées particulièrement importantes puisqu'elles sont sollicitées plus tard et conditionnent les habiletés de compréhension en lecture.

3. Les principaux modèles de l'apprentissage de la lecture

Dans la littérature relative à l'apprentissage de la lecture, trois types de modèles permettent de rendre compte de l'identification des mots écrits : les modèles développementaux, les modèles par analogie et les modèles interactifs. Les premiers décrivent, en termes de développement, les différentes étapes par lesquelles passe l'enfant lors de ses premiers contacts avec l'écrit. Aux modèles développementaux, s'opposent depuis quelques années des modèles dits

fonctionnels ou interactifs insistant sur l'activation simultanée de processus orthographiques et phonologiques. Dans ces modèles, les connaissances orthographiques élaborées à partir du repérage de leurs correspondants phonologiques seraient donc utilisées plus précocement que dans les modèles développementaux. Enfin, il existe d'autres modèles de l'apprentissage de la lecture, comme celui de Goswami & Bryant (1990), qui accordent une place importante à l'utilisation des processus analogiques. Tout d'abord, nous présenterons les modèles en stades, adaptation développementale des modèles à deux voies, puis ceux mettant en avant le rôle des analogies dans la lecture. Enfin, seront présentés les modèles interactifs qui deviennent actuellement les modèles les plus influents dans le domaine de la lecture.

3.1. Les modèles développementaux en stades

La plupart des modèles développementaux de la lecture suggèrent que les lecteurs débutants passent à travers une série d'étapes, chacune marquée par l'adoption de stratégies ou de procédures de traitement différentes conduisant à la mise en place progressive des deux voies qui caractérisent la lecture experte (Frith, 1985). Dans ces modèles, l'ordre des étapes de succession est strict et identique pour tous les enfants. Le passage d'une étape à l'étape suivante s'effectue sous l'effet de l'instruction et ne peut avoir lieu que si les compétences propres à l'étape antérieure sont parfaitement maîtrisées. Ainsi, chaque étape suppose un fonctionnement spécifique qui implique une homogénéité dans les conduites des enfants[4]. Le modèle de Frith (1985) reconnait l'existence de trois étapes ou procédures : la procédure logographique qui consiste à utiliser des indices visuels, la procédure alphabétique qui se traduit par un recours systématique à la médiation phonologique et enfin la procédure orthographique, celle de l'expertise en lecture.

La procédure logographique qui domine la première étape de la lecture correspond à un traitement visuel des mots sans recours à la phonologie et à l'ordre des mots. Les sujets utilisent des indices fournis par le contexte extra-linguistique dans lequel se trouvent les mots (le logo rouge de Coca Cola par exemple). A ce niveau, les enfants ont également recours pour reconnaître le mot à la présence de traits visuels saillants, notamment certaines lettres ou configuration de lettres. Ces différentes modalités de traitement de l'information permettent la constitution d'un premier vocabulaire global. Selon Seymour & Elder (1986), un lexique d'une centaine de mots peut être ainsi discriminé. Pour se faire, les enfants peuvent s'appuyer

[4] Cette position est fortement critiquée par Seymour (1989) qui affirme que les lecteurs débutants se caractérisent par une forte hétérogénéité interindividuelle des comportements en partie liée aux méthodes d'enseignement.

sur différents éléments : la longueur des mots (court, moyen, long), la position dans le mot de l'indice connu (gauche, centre, droite) et sur les quelques formes saillantes qu'ils connaissent. Au cours de cette période, l'enfant mémorise des formes visuelles et ne dispose pas de procédures de traitement spécifiques au langage. Actuellement, la plupart des chercheurs ont plutôt tendance à ne plus considérer ce type de procédure comme une étape dans le cursus développemental. Selon Sprenger-Charolles & Colé (2003), « *les prélecteurs, à cette étape de leur développement lexique, semblent lire plus l'environnement que le mot lui-même* » (p. 77).

C'est au cours de la seconde étape que les enfants vont développer les premières procédures d'analyse linguistique. En effet, la procédure alphabétique est caractérisée par le rôle premier des facteurs phonologiques et, contrairement à l'étape précédente, l'identité et l'ordre de toutes les lettres sont déterminants. Les enfants apprennent à ce niveau à mettre en correspondance l'écrit et l'oral. La reconnaissance des mots se fait par recours systématique à des règles de correspondance entre des unités visuelles discrètes (les graphèmes) et des unités phonémiques abstraites (les phonèmes). Il s'agit d'une procédure analytique dans laquelle tous les graphèmes sont traités, décodés les uns après les autres. La procédure phonologique, systématiquement utilisée par les enfants au début de l'apprentissage de la lecture, s'apparente à la voie indirecte ou voie d'assemblage phonologique décrite dans le modèle à double voie. L'utilisation de cette voie indirecte va permettre aux enfants de se construire un lexique mental.

Une fois le système de conversion phonologique maîtrisé, une procédure de reconnaissance plus rapide se met en place. Au cours de cette dernière procédure, appelée procédure orthographique, les représentations lexicales des mots deviennent accessibles. L'existence d'un lexique mental permet la mise en correspondance directe des mots déjà rencontrés avec leurs correspondants stockés dans le lexique mental. A partir de ce moment, la reconnaissance des mots écrits ne nécessite plus, comme dans le stade alphabétique, un recours systématique à la conversion phonologique. Cette procédure s'apparente à la voie directe, par adressage, du modèle à double voie. Dès la fin de la première année d'apprentissage de la lecture, certains mots (les plus fréquents) sont ainsi reconnus automatiquement. L'enfant a stocké en mémoire à long terme des informations spécifiques sur les caractéristiques orthographiques des différents mots de la langue rencontrés jusque-là. De cette façon, il va retrouver directement les mots en mémoire sur la base de leurs segments orthographiques. En effet, les enfants disposent maintenant, comme les lecteurs experts, des deux procédures d'identification des mots : l'assemblage et l'adressage. Au cours de cette étape, l'identification des mots écrits

repose de plus en plus sur des processus automatiques. En conséquence, son coût attentionnel diminue, ce qui permet au lecteur d'allouer davantage de ressources cognitives aux aspects sémantiques et syntaxiques de la phrase (Demont & Gombert, 2004).

Les premiers travaux sur l'évolution des stratégies d'identification des mots écrits en arabe (Ammar, 1997 ; Taouk & Coltheart, 2004) ont pu montrer que les stratégies de type logographique, alphabétique et orthographique décrites à propos de la langue anglaise sont utilisées par les lecteurs en arabe. En effet, les résultats obtenus par Ammar (1997) montrent que les élèves bons lecteurs de 1ère année de l'école primaire mobilisent des stratégies de recodage phonologique à travers la conversion du symbole écrit en son. A partir de la 3ème année, les bons lecteurs ont atteint la phase ultime du développement de la compétence en lecture et sont capables de mettre en œuvre des stratégies de type orthographique. En revanche, les faibles lecteurs de 1ère année ont recours à des stratégies de type logographique en manifestant une sensibilité à la proximité visuelle. Les résultats obtenus par les élèves faibles lecteurs scolarisés en 2ème et 4ème année suggèrent un traitement à la fois phonologique et visuel au cours de l'identification des mots.

Les modèles développementaux constituent toujours à l'heure actuelle des modèles de référence et ont servi de cadre à de multiples recherches dans le domaine de l'apprentissage de la lecture. Un intérêt majeur de ce type de modèle est d'avoir mis l'accent sur l'importance des correspondances grapho-phonologiques (CGP) lors des étapes initiales de l'apprentissage. Ces modèles sont actuellement remplacés par des modèles plus dynamiques qui décrivent comment le système de traitement de l'information mobilise différents processeurs dans des tâches de reconnaissance des mots écrits. Certains modèles accordent une place importante à l'utilisation des connaissances antérieures (phonologiques et visuo-orthographiques) qui résulte de la mise en œuvre de processus analogiques.

3.2. Les modèles de lecture par analogie

Dans les modèles de lecture par analogie, la nature non lexicale de la procédure indirecte est mise en cause. Ces modèles insistent fortement sur l'ancrage de nouvelles connaissances antérieures et plus spécifiquement sur les liens de niveau phonologique et orthographique. Treiman (1989) a montré que les correspondances entre l'écrit et l'oral ne sont pas restreintes au niveau des simples phonèmes, mais que les lecteurs effectuent également des appariements qui se situent au niveau des unités infra-syllabiques. On distingue alors un niveau intermédiaire entre la syllabe et le phonème constitué par les éléments infra-syllabiques. En

effet, la syllabe est constituée de deux unités principales : l'attaque constituée du groupe consonantique initial (par exemple « tr » dans « trois ») et la rime constituée par la voyelle et les éventuelles consonnes qui la suivent. Dans l'apprentissage de la lecture, l'enfant peut utiliser ce type d'unité pour lire les mots par analogie. Ce type de procédure revient à pouvoir lire un mot nouveau (mot cible) à partir d'un mot connu (mot source ou mot indice).

Selon le modèle de Goswami & Bryant (1990), la lecture par analogie est plus précoce que la lecture par application des règles de correspondances graphèmes-phonèmes puisque la conscience des rimes sur laquelle la première s'appuie est plus précoce que la conscience phonémique sur laquelle la seconde s'appuie. Ainsi, les débuts de la lecture ne sont pas visuels mais basés sur des unités phonologiques (attaque-rime). Dans une première étude, Goswami (1986) a montré que les prélecteurs et les apprentis lecteurs utilisent précocement l'analogie pour lire des mots nouveaux. Elle propose à des enfants de 5 à 8 ans une tâche de lecture de mots suivant ou non un mot indice (exemple *beak).* Les mots peuvent être analogues à l'indice (*pean, peak)* ou non (*bask, lake*). Trois groupes de niveau en lecture furent définis à partir d'un test de lecture de mots, les non lecteurs (âge lexique : 6 ans ; âge chronologique : 5,4 ans) et de groupes de lecteurs de niveaux lexiques différents, L1 (âge lexique : 6,10 ans ; âge chronologique : 6,10 ans) et L2 (AL : 7,4 ans ; AC : 7,1 ans). Les résultats montrent que les enfants lisent plus de mots quand l'analogie peut être faite à partir de l'indice que quand aucune base pour l'analogie n'est donnée. D'autre part, il apparaît que les trois groupes d'enfants peuvent faire des analogies entre les fins des mots (*peak* par *beak*) mais seuls les groupes 2 et 3, plus avancés, peuvent en faire sur les débuts de mots.

Des résultats similaires ont été retrouvés auprès de jeunes enfants français, élèves de CP (Faure, Ecalle, Magnan, 2000). La tâche consistait à présenter successivement deux mots associés chacun à une image. L'enfant devait indiquer si le second mot, après lecture du premier par l'expérimentateur, allait bien avec son image. Dans la condition « analogie », les mots partageaient une unité large en position initiale (*but, bus*), en final (*fer, mer*) et une unité réduite en initial (*bol, bec*) et en final (tir, mur). Dans la condition « sans analogie », l'item indice était un pseudo-mot qui représentait un animal imaginaire. Les auteurs observent une réussite significativement supérieure pour les items partageant la rime (unité large en position finale) dans la condition « analogie » alors qu'il n'y a aucun effet dans la condition « sans analogie ».

Le caractère précoce de la lecture par analogie est contesté par différents auteurs, notamment les partisans des modèles en stades. Selon cette conception, un minimum de connaissance des CGP est nécessaire pour utiliser les analogies, Ehri & Robbins (1992) comparent les résultats de deux groupes de lecteurs : un groupe de « décodeur » et un groupe de « non décodeurs » sélectionnés en fonction de leurs scores en lecture de pseudo-mots. Dans la condition analogie, les sujets apprennent des items analogues à ceux qui seront ensuite utilisés comme items test. Ces analogies concernent la rime (exemple SAAV *vs* KAAV). Dans la condition contrôle, les sujets apprennent des items de même structure graphémique mais qui ne sont pas analogues aux items test (exemple SAAV *vs* RAAN). Les résultats montrent que seuls les décodeurs peuvent lire les mots par analogie. Les non décodeurs lisent les mots-test en se basant sur les lettres communes. Toutefois, parmi les décodeurs, les sujets de la condition « analogie » lisent plus de mots que ceux de la condition « contrôle ». Ces données indiquent, d'une part, qu'il est nécessaire de posséder un minimum d'habileté de décodage phonologique pour lire les mots par analogie et, d'autre part, que la fusion d'unités de type attaque-rime est plus facile que la fusion d'unités phonémiques[5].

Le caractère tardif du recours à l'analogie a aussi été mis en évidence dans des tâches de lecture à voix haute. Selon Casalis (1995), si la lecture par analogie a lieu avant le traitement alphabétique, les lecteurs les moins avancés (CP et dyslexiques) devraient lire davantage de pseudo-mots analogues que de pseudo-mots non analogues. L'auteur a proposé une tâche de lecture à voix haute de pseudo-mots analogues de différentes longueurs (*nobe / robe*, *pardin / jardin*) et de pseudo-mots contrôle. Quatre groupes de sujets effectuaient cette tâche pilotée par ordinateur, des dyslexiques et des normo-lecteurs appariés sur le niveau d'âge lexique (CP et CE1). Les résultats indiquent que l'analogie n'a pas d'effet global. Cependant, les comparaisons menées sur chacun des sous-groupes font apparaître que l'analogie n'intervient que sur les temps de lecture des normo-lecteurs de CE1. Il semblerait donc que seuls les lecteurs les plus avancés exploitent le caractère analogique des items puisqu'ils sont lus plus rapidement.

Selon Ecalle & Magnan (2002), une façon de sortir de ce débat sur l'apparition précoce ou tardive de l'analogie en lecture est de « *considérer qu'au tout début de l'apprentissage de la*

[5] Une différence importante entre cette étude et celles de Goswami est qu'elle prend en compte le fait que la lecture par analogie fait appel à la mémoire alors que dans la plupart des expériences de Goswami, les sujets ont des mots indices sous les yeux.

lecture, l'analogie consisterait à activer des liens entre des unités phonologiques prégnantes (la rime, la syllabe) et des unités orthographiques stables » (p. 56). Dans ce cas, les procédures analogiques participeraient à la catégorisation d'unités linguistiques à partir desquelles s'élaborent progressivement les CGP. L'interactivité entre les différents processus est particulièrement bien représentée dans les modèles interactifs pour lesquels la connaissance d'un mot suppose que toutes les connaissances associées à ce mot (phonologiques, orthographiques, sémantiques et morphologiques) sont utilisées par le système.

3.3. Les modèles connexionnistes et interactifs

L'approche traditionnelle des modèles à double voie (Coltheart, 1978) considère que les connaissances des relations lettres-sons peuvent être exprimées sous la forme d'un ensemble restreint de règles de correspondances graphèmes-phonèmes (CGP). Le recours à ce système permet de lire les mots réguliers et de prononcer les mots nouveaux et les pseudo-mots. Cette procédure, appelée voie indirecte, ne peut en revanche, être utilisée pour lire les mots irréguliers qui reposent sur le traitement d'un code orthographique (procédure d'adressage ou voie directe). Dans ce cas, la prononciation du mot s'effectue sur la base d'un code phonologique récupéré directement dans le lexique mental. Dans cette perspective, les informations orthographiques et phonologiques apparaissent comme des sources d'activation possibles et indépendantes qui opèrent en parallèle. Actuellement, les résultats des recherches ont conduit à proposer d'autres types de modèles plus dynamiques tels que les modèles connexionnistes qui se distinguent des modèles à double voie dans la manière d'envisager le format de stockage des connaissances linguistiques manipulées pendant la lecture et dans le type de procédures mises en œuvre par le lecture sur ces connaissances.

3.3.1. Le modèle connexionniste de Seidenberg & Mc Clelland (1989)

Dans une perspective connexionniste, un système cognitif est un ensemble 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. Une stimulation déclenche l'activation de ce système jusqu'à l'émergence d'un patron stable d'unités activées qui constituera la réponse du réseau. En ce qui concerne l'apprentissage de la lecture, l'un des modèles de référence est celui de Seidenberg & Mc Clelland (1989). Dans ce modèle, les réseaux impliquent des connaissances orthographiques, phonologiques et sémantiques qui interagissent ainsi que des processus de

transformation permettant la transmission d'excitations ou d'inhibitions entre les unités de connaissance. Chaque unité est caractérisée par son niveau d'activité mais, à chaque instant, les unités modifient leur activation en fonction des signaux qu'elles reçoivent. Les codes orthographiques et phonologiques sont représentés par des patterns d'activation des unités. Ce système ne comprend donc pas de règles de CGP et il n'y a plus à proprement parler de lexique mental, dans le sens où il n'y a pas de mots stockés comme des unités en mémoire à long terme. La reconnaissance des mots est le produit d'une distribution d'une activation dans le système de traitement des lettres et des informations phonologiques et sémantiques associées à ces lettres.

Les connexionnistes réfutent l'idée selon laquelle la lecture de mots fréquents se ferait en consultant le lexique orthographique alors que celle des mots rares et des pseudo-mots nécessiterait l'utilisation des règles de correspondances graphophonologiques. Il n'y a donc pas, contrairement aux modèles à double voie, de distinction entre la lecture des mots réguliers et irréguliers. Un mécanisme unique permet de rendre compte de la lecture de tous les items quelle que soit leur nature. Ainsi, pour les mots fréquents, qui sont présentés plus souvent au réseau, les résultats ne montrent pas de différence en fonction de la régularité. La distinction entre mots réguliers et irréguliers se trouvant éliminée par le nombre important de présentation. En revanche, pour les mots rares, l'effet de la régularité persiste et leur prononciation dépend de leurs propriétés internes. Ces deux effets (fréquence et régularité) sont la conséquence d'un ajustement plus fréquent pour les items fréquents ou réguliers, que pour les items rares ou irréguliers.

Si le réseau stimule correctement les performances du lecteur adulte pour les mots, il n'en est pas de même pour les pseudo-mots. Les connexions avec d'autres mots sont moins nombreuses pour un pseudo-mot que pour un mot dans la mesure où, par définition, un pseudo-mot correspond à une structure orthographique qui n'existe pas dans la langue. Ainsi, le pseudo-mot « nirpe » active moins de connexions potentielles qu'un mot de même structure orthographique tel que « porte ». Les effets de fréquence et de régularité sont la conséquence d'un ajustement plus fréquent pour les items fréquents ou réguliers, que sur les items rares ou irréguliers. Ils sont donc issus d'une même source.

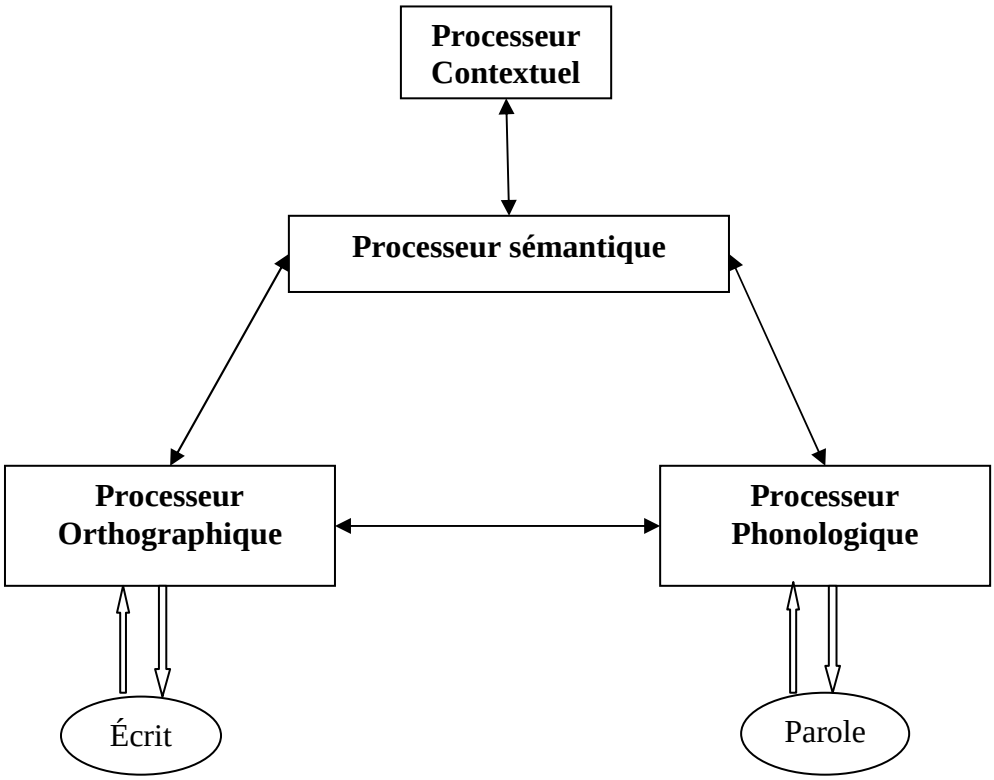


Figure 1 : Modèle connexionniste adapté de Seidenberg & McClelland (1989)

Le principe fondamental des modèles connexionnistes est qu'un mécanisme unique permet de rendre compte de la lecture de tous les mots quelles que soient leur fréquence et leur régularité. Ce mécanisme unique permet également de rendre compte de la lecture des pseudo-mots. Les mots familiers bénéficient de la rapidité d'activation consécutive aux lectures antérieures. Ceci est en contradiction avec les modèles à double voie, selon lesquels, la lecture des mots fréquents se ferait en consultant le lexique interne par voie directe, alors que celle des mots rares et des pseudo-mots nécessiterait l'utilisation de règles de CGP. L'inconvénient majeur de ces modèles est qu'ils ne prennent pas en compte les connaissances métaphonologiques, ni les relations entre ces connaissances et les mécanismes de lecture. Or, il a été montré que les capacités métaphonologiques jouent un rôle primordial dans cette acquisition.

Ces modèles prévoient que, dès le départ, il existe des unités spécialisées dans le traitement de l'orthographe (unités orthographiques). Il est toutefois possible, en gardant la même perspective de décrire l'apparition des unités orthographiques dans un système qui en est initialement dépourvu. Il suffit, pour cela, d'envisager la spécialisation d'unités qui,

initialement réagissent à des percepts visuels non linguistiques. C'est l'idée défendue par Gombert, Bryant & Warrick (1997).

3.3.2. Le modèle de Gombert, Bryant & Warrick (1997)

S'inspirant de la perspective connexionniste proposée par Seidenberg & Mc Clelland (1989) et s'appuyant sur le modèle développemental de Frith (1985), le modèle de Gombert, Bryant & Warrick (1997) vise à décrire comment le système de traitement de l'information mobilise les différentes connaissances associées au mot (phonologique, orthographique, sémantique et morphologique) dans l'identification des mots écrits et comment il devient progressivement capable de traiter le langage écrit sous l'effet de l'enseignement et de l'expérience avec l'écrit. Ce modèle cherche à rendre compte, d'une part, de l'articulation entre le traitement du langage oral et le traitement du langage écrit et, d'autre part, du rôle des connaissances implicites et explicites dans l'apprentissage de la lecture (*cf.* figure 2). Deux types de connaissances sont ainsi à l'œuvre dans la lecture, des connaissances implicites et opérationnelles qui interviennent à l'insu du lecteur et des connaissances explicites dont il contrôle consciemment l'intervention (Gombert, Bonjour & Marec-Breton, 2004).

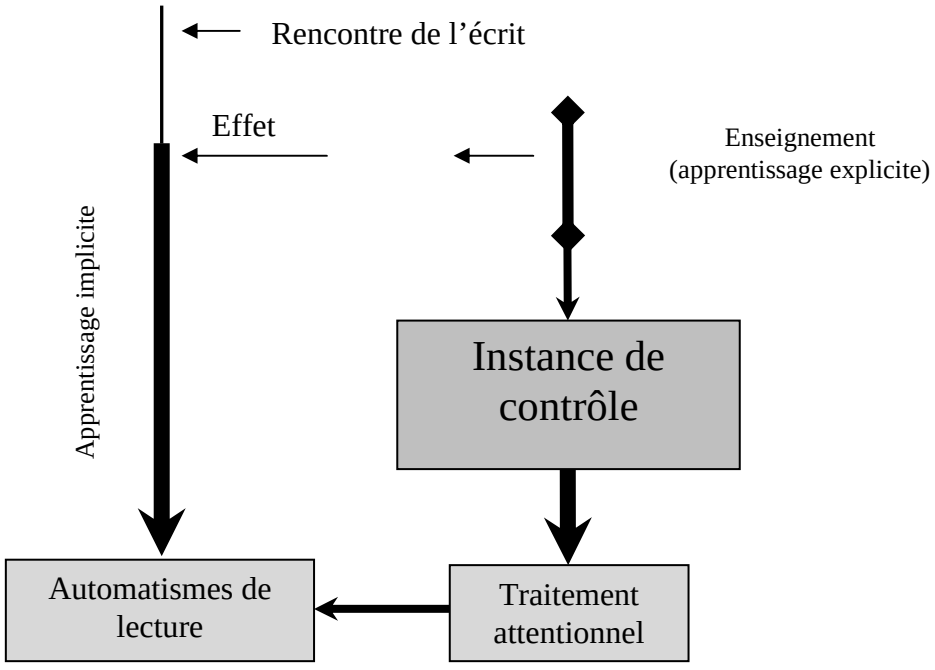


Figure 2 : Apprentissage implicite et explicite de la lecture (d'après Gombert, 2003)

Les modèles développementaux font, en effet, peu de place à la description de la base de connaissance préexistant à l'apprentissage et se centrent essentiellement sur les changements

intervenant sous l'effet de l'enseignement. Or, selon Gombert (2003), avant même l'apprentissage de la lecture, l'enfant dispose d'un système capable de traiter le langage oral qui servira de base à l'élaboration du système de traitement de l'écrit. Ce système comprend différents processus de traitement de l'information langagière non spécifique à la lecture : un *processeur pictural* traiterait l'information visuelle, un *processeur phonologique* traiterait l'information auditive, un *processeur sémantique* assurerait l'attribution de la signification des mots et des relations sémantiques existant entre eux et enfin, un *processeur contextuel* élaborerait une interprétation cohérente de ces informations.

L'enfant reconnaît, très tôt, un certain nombre de mots d'une manière logographique par mémorisation d'indices visuels saillants. Le processeur pictural commence alors à traiter l'information écrite de manière spécifique en se spécialisant dans le repérage des configurations orthographiques des mots. Les informations lexicales traitées conjointement par les processeurs pictural, phonologique et sémantique sont mises en relation et aboutissent à la création d'un processeur orthographique spécialisé dans le traitement des lettres. L'enfant peut utiliser les connaissances visuo-orthographiques, phonologiques et morphologiques acquises implicitement pour reconnaître des mots écrits et pour prononcer des configurations orthographiques jamais rencontrées jusqu'alors, ce qui conduit à l'accroissement de son lexique écrit. De la même façon, dès ses premières tentatives de lecture, le débutant peut faire des analogies entre les mots qu'il sait déjà reconnaître et ceux qu'il tente de découvrir. Le terme de traitement analogique a été proposé pour désigner le fait que l'enfant utilise ses connaissances sur la prononciation de certaines configurations orthographiques pour traiter les mots nouveaux. Ces connaissances rendent l'enfant capable de faire des inférences au sujet des mots non familiers. Dans un premier temps, il semble que l'enfant lecteur n'a pas conscience des analogies qu'il opère. Progressivement, l'activation répétée de ces connaissances acquises implicitement renforcera les relations entre des configurations orthographiques et les configurations phonologiques et morphologiques correspondantes et de ce fait, augmentera leur disponibilité pour l'accès conscient. C'est la confrontation du système avec l'écrit qui provoquera cette évolution.

D'après ce modèle, deux processeurs, visuel et phonologique, interviennent conjointement. Le premier, en établissant une base de connaissances visuelles, est capable de traiter globalement des mots et certaines configurations orthographiques fréquentes. Le second participe à la constitution d'une base de connaissances phonologiques parmi lesquelles la sensibilité à la rime. La mise en œuvre précoce des procédures analogiques (souvent implicites) participe

également au développement et à la maîtrise du code alphabétique (de nature explicite) puis orthographique via l'enseignement. Cependant, les traitements implicites ne s'avèrent pas suffisants pour permettre une lecture habile et ne permettent pas à eux seuls la connaissance des règles à appliquer pour lire n'importe quel mot dans n'importe quel contexte. C'est toutefois sur la base des apprentissages implicites que l'instruction explicite de la lecture permettra l'installation des traitements alphabétiques et orthographiques indispensables pour une lecture habile. Ainsi, une bonne compréhension de l'ensemble des processus à l'œuvre dans l'apprentissage de la lecture exige que l'on considère de façon complémentaire les versants implicites et explicites de cet apprentissage (Gombert, 2003).

4. Spécificités de la lecture en langues sémitiques

La plupart des travaux ayant traité de l'identification des mots écrits se sont intéressés à l'anglais. Les études sur les langues sémitiques, comme l'arabe et l'hébreu, sont très rares. Ces deux langues utilisent une orientation droite-gauche et possèdent les mêmes caractéristiques morphologiques[6]. La majorité des mots sont formés par l'association d'une racine et d'un pattern vocalique. Dans les deux langues les racines sont trilitères c'est à dire composée de trois consonnes et les dérivations sont obtenues par l'ajout d'affixes et par flexion interne. Une autre particularité est liée au fait que certaines lettres changent de forme selon leur position dans le mot. En arabe, vingt deux des vingt huit lettres changent de formes selon leur position (initiale, médiane, finale, isolée) dans le mot (par exemple, le phonème /h/ [هـ] /ha : ʔ/ est représenté par différents graphèmes : isolée [ه], au début du mot [هـ], au milieu [ﻫـ], et à la fin [ـه]). En hébreu, il y n'a que cinq lettres qui changent de formes en position finale (מ-ם, כ- ך,פ-ף,צ-ץ,נ- ן).

Contrairement aux langues latines où les voyelles font partie de l'alphabet et sont représentées par des lettres, pour les langues sémitiques, les voyelles ne font pas partie de l'alphabet, elles sont représentées par des signes diacritiques ajoutés aux mots. Ainsi, l'arabe et l'hébreu s'écrivent toutes les deux sous une double forme, vocalisée et non vocalisée. La forme non vocalisée, qui constitue l'écriture standard, note uniquement le squelette du mot composé de consonnes et de voyelles longues[7]. Lorsque l'écriture est vocalisée, les voyelles sont ajoutées par des signes diacritiques situés au-dessous et au-dessus des consonnes (*cf.* tableau 1). Dans l'écriture vocalisée, les correspondances graphophonologiques sont essentiellement

[6] Une description détaillée des caractéristiques de la langue arabe écrite sera présentée dans le chapitre suivant.
[7] En Tunisie, l'écriture non vocalisée est imposée progressivement dès la fin de l'école primaire.

régulières, elle est donc généralement destinée aux apprentis lecteurs. Lorsque la vocalisation est absente, la prononciation et la signification des mots, qui se comportent comme des homographes, ne peuvent être obtenues que par référence au contexte de la phrase. Ces données laissent supposer que la présence de la vocalisation permet une identification plus rapide à partir de l'information orthographique. Dès lors, si le lecteur recourt à la procédure d'assemblage pour identifier les mots, les mots vocalisés seront lus plus rapidement que les mots non vocalisés.

Tableau 1: Exemple en arabe d'une forme non vocalisée à laquelle peuvent correspondre trois mots de significations différentes

Mot écrit	**Mot arabe**	**Prononciation**	**Traduction**
/k.t.b/ كتب	كَتَبَ	/Kataba/	Il a écrit
	كُتِبَ	/kutiba/	Il a été écrit
	كُتُبٌ	/kutubun/	Livres

Les données recueillies en décision lexicale suggèrent que l'information apportée par les signes diacritiques est peu utilisée dans l'identification des mots. L'étude de Bentin & Frost (1987) montre que la présence des voyelles retarde les décisions lexicales, que les mots soient fréquents ou rares, avec des racines ambigües ou non. Ils ont pu observer que les mots pointés étaient prononcés plus rapidement que les mots non pointés. L'information apportée par les signes diacritiques est peu exploitée lorsqu'elle n'est pas nécessaire pour distinguer entre les mots et les pseudo-mots. Koriat (1985) montre que la présence des marques diacritiques facilite plus la décision lexicale pour les mots rares que pour les mots fréquents. Cette observation est compatible avec l'idée d'une médiation phonologique plus fréquente pour l'identification des mots rares que pour les mots fréquents. Frost & Bentin (1992) font l'hypothèse que les lecteurs ont développé un lexique orthographique basé sur des séquences non vocalisés servant d'interface à la phonologie de plusieurs mots. Ce rôle de bruitage de la vocalisation qui perturbe ou ralentit l'identification des mots a déjà été mis en évidence par Roman (1982) et plus récemment par Ammar (1997) sur la langue arabe. Ammar (1997) a montré, à partir d'une épreuve de détection d'erreurs orthographiques sur mots vocalisés présentés en contexte imagé, une négligence des signes vocaliques. Les perturbations sur les consonnes étaient beaucoup mieux détectées que les perturbations vocaliques et ceci, aussi bien par les bons lecteurs que par les faibles lecteurs.

Abu Rabia (1998) a étudié l'effet de la vocalisation et du contexte en utilisant différents types de matériel : poèmes, textes coraniques, articles de journaux et des contes. Les sujets avaient pour tâche de lire des listes de mots et de paragraphes en version vocalisée et non vocalisée. Un effet significatif de la vocalisation et du contexte a été observé aussi bien chez les bons lecteurs que les faibles lecteurs et pour chaque type de texte. Dans une autre expérience, Abu Rabia (2001) avait pour objectif d'explorer l'effet de la présence ou de l'absence de voyelles sur les processus de lecture en arabe et en hébreu chez des sujets adultes ayant l'arabe pour langue maternelle et l'hébreu comme une langue seconde. Les sujets (N=65) avaient pour tâches de lire des listes entières de mots et de paragraphes en arabe vocalisé et non vocalisé ainsi que des listes en hébreu pointé et non pointé. Ils devaient lire également deux histoires, l'une en arabe (vocalisé et non vocalisé) et l'autre en hébreu (pointé et non pointé). Les résultats révèlent un effet significatif de la vocalisation dans toutes les conditions de l'expérience en arabe et en hébreu. Les voyelles en arabe et en hébreu apportent une information phonologique importante au mot. De la même façon, il s'avère que la compréhension est meilleure dans le cas d'un texte entièrement vocalisé par rapport à un texte non vocalisé. Ainsi, l'information phonologique supplémentaire apportée par les voyelles est, selon l'auteur, importante pour la compréhension d'un texte en arabe même pour des lecteurs confirmés.

Dans une expérience plus récente, Ammar (2003) a administré à des élèves de 2ème, 4ème et 6ème année une épreuve de lecture à voix haute de pseudo-mots. Il constate que l'information vocalique procurée par les signes diacritiques n'est pas ignorée dans la lecture à haute voix. Elle est surtout utilisée dans la prononciation des pseudo-mots dont la perturbation porte sur les consonnes ou sur les voyelles. Conformément au modèle à double voie, la procédure d'adressage phonologique serait utilisée dans la prononciation des mots non vocalisé et des mots vocalisés fréquents. En revanche, le lecteur utiliserait la procédure d'assemblage phonologique pour prononcer les mots vocalisés rares. Ainsi, la vocalisation n'est pas totalement ignorée pendant la lecture ; elle est d'autant plus utilisée que le mot est rare.

Bien que l'orthographe arabe dans sa forme vocalisée se caractérise par des correspondances graphophonologiques (CGP) strictement régulières, Ammar (2003) pense que le lecteur utilise aussi des analogies orthographiques. La structure morphologique particulière de la langue arabe fondée sur la racine consonantique laisse supposer que les mots sont plus facilement décomposables en racine consonantique et morphèmes vocaliques qu'en attaque et rime. Les

analogies orthographiques sont établies par récupération en mémoire de tous les mots qui partagent avec l'item cible la même séquence consonantique correspondant à la racine.

5. Conclusion

Au cours de ce chapitre introductif, nous avons insisté sur l'importance de la reconnaissance des mots écrits comme base indispensable pour accéder à la lecture. Il parait en effet inconcevable d'accéder à la signification du texte sans pouvoir identifier au moins une partie des mots qu'il comporte. L'identification doit être non seulement correctement effectuée mais également doit être au moins partiellement automatisée. L'automatisation progressive des processus d'identification des mots écrits libère les ressources cognitives nécessaires à la compréhension. Chez le lecteur habile, les activités de d'identification et de construction de la compréhension se déroulent simultanément et en interaction. Apprendre à lire consiste ainsi à développer et automatiser les composantes spécifiques à la lecture.

Différentes modélisations ont été proposées pour décrire comment l'enfant parvient à la maîtrise et l'automatisation des procédures d'identification des mots écrits. Les modèles développementaux à stades consistent à décrire des changements qualitatifs dans les procédures successivement élaborées par l'apprenti lecteur, d'autres modèles mettent pour leur part l'accent sur l'interaction des différentes composantes de l'apprentissage de la lecture. Un intérêt majeur des modèles développementaux est d'avoir mis en exergue le rôle fondamental des connaissances phonologiques lors des premières étapes de l'apprentissage de la lecture. En effet, la compréhension du principe alphabétique nécessite de la part de l'enfant des connaissances explicites sur la structure phonologique du langage. Ces connaissances explicites s'avèrent ainsi déterminantes pour permettre au jeune lecteur d'acquérir et d'utiliser le décodage phonologique nécessaire à la lecture des mots qu'il n'a encore jamais rencontrés sous forme écrite.

En revanche, cette prédominance des modèles développementaux commence à décliner principalement parce qu'ils s'avèrent incapables d'expliquer de façon satisfaisante un apprentissage dont la dynamique doit prendre en considération le fait que certaines procédures peuvent coexister et ce, dès le début de l'apprentissage. Il est à souligner, d'autre part, le fait que ces modèles et leurs évaluations sont essentiellement anglophones. Cette situation est dommageable dans la mesure où la trajectoire développementale peut être influencée par certains principes généraux communs à toutes les langues mais également par des variations liées aux spécificités de chacune. Il est en particulier primordial d'évaluer les hypothèses

développementales présentées en travaillant sur d'autres langues que l'anglais, notamment sur les langues sémitiques.

Nous allons exposer plus en détails les caractéristiques de la langue arabe et les spécificités de l'apprentissage de la lecture chez le lecteur arabophone dans le chapitre suivant.

Chapitre 2 : La langue arabe : spécificités et apprentissage

1. Historique

La langue arabe, dans sa forme écrite, est l'une des plus anciennes langues vivantes dans le monde actuel ; et ce, tout en continuant à développer des niveaux dialectaux ou intermédiaires.[8] Avant l'avènement de l'islam, l'arabe était en contact avec les langues appartenant à la famille sémitique comme le phénicien, l'hébreu, l'araméen, etc. Au sein de cet ensemble, elle appartient au sous groupe du sémitique méridional[9]. Avec l'extension de l'empire arabo-islamique à partir du VII^{ème} siècle, l'arabe est devenu l'une des principales langues véhiculaires du monde. Elle a connu, au cours de sa longue histoire, d'importantes périodes de création littéraire. C'est grâce à l'immense apport coranique et à toute la production religieuse qu'il a engendrée que l'arabe dit « classique » a succédé à l'arabe ancien. Cette mutation est attestée en particulier par le développement de l'écriture, enrichie par les signes diacritiques et vocaliques à valeur graphémo-phonologique et par la codification de la langue et l'apparition des premiers dictionnaires et traités de grammaire. Ces diverses influences sont à l'origine des variétés d'arabe qui ont évoluées vers les parlers actuels pratiqués dans chacun des pays arabes.

L'évolution de l'arabe classique et de ses dialectes pendant un millénaire finit par engendrer ce qu'on appelle « l'arabe moderne », langue de la presse écrite et audio-visuelle, des débats politiques, des textes scientifiques et de plus en plus celle des textes littéraires. Parlé dans la plupart des pays arabes, l'arabe moderne n'est en revanche presque jamais la langue des échanges quotidiens. Les diverses formes d'arabe dialectal diffèrent les unes des autres et de l'arabe moderne sur le plan de la prononciation, du vocabulaire et de la grammaire. Toutefois, l'influence stylistique du français et de l'anglais est fréquente dans de nombreux cas. Au XX^{ème} siècle, en particulier, une grande partie du vocabulaire scientifique a été emprunté au français et à l'anglais.

[8] Aujourd'hui, l'arabe est la langue officielle de 22 pays (250 millions d'habitants) et pratiqué également dans plusieurs pays musulmans non arabes comme la Turquie ou l'Iran dont les langues ont connu, au cours de l'histoire, diverses formes d'interférence.

[9] Les langues sémitiques font partie de la famille des langues afro-asiatiques, et sont parlées en Afrique ainsi qu'au Proche et Moyen-Orient.

A l'époque moderne, l'arabe peut être considéré comme un terme générique regroupant plusieurs variétés :

- l'arabe classique : la langue du Coran, parlée au VII^{ème} siècle.
- l'arabe standard moderne : une forme un peu différenciée de l'arabe classique qui constitue la langue écrite. Il s'agit de la variété retenue comme langue officielle dans tous les pays arabes et comme langue commune entre eux. Elle est également la langue employée dans la plupart des écrits administratifs, médiatiques, scientifiques, techniques, littéraires ainsi que dans la majorité des journaux télévisés.
- l'arabe dialectal : langue parlée au quotidien dans les pays arabes. D'un pays à l'autre, voire d'une région à l'autre dans un même pays, les dialectes arabes varient parfois tellement qu'il est difficile de les considérer comme une seule langue (Baccouche, 2003).

2. Caractéristiques de la langue arabe

2.1. Système d'écriture

L'arabe est une langue sémitique qui s'écrit et se lit de droite à gauche. L'alphabet arabe comporte 28 graphèmes consonantiques dont deux semi-consonnes ([و] /w/ et [ي] /j/) et de 6 graphèmes vocaliques (trois voyelles courtes : [ـُ] /u/ , [ـَ] /a/ et [ـِ] /i/ et trois voyelles longues : [ـُو] / u : / , [ـَا] / a : / et [ـِي] / i : /). Les lettres *alif* [ا], *wa:w* [و] et *ja:* [ي] notent tantôt des consonnes, tantôt des voyelles longues. En arabe, seules les consonnes s'écrivent obligatoirement, l'adjonction des voyelles étant facultative[10].

La plupart des lettres de l'alphabet arabe changent de forme suivant qu'elles sont isolées, au début, au milieu ou à la fin du mot (voir tableau 2). En arabe, 22 lettres parmi les 28 de l'alphabet ont chacune quatre formes (i.e., le phonème [ه] /h/ est représenté par différents graphèmes : [ﻫ], [ﻩ], [ﻬ], et [ﻪ]) et 6 lettres ont deux formes chacune. La distinction entre certaines lettres se fait sur la base de l'existence, la place et le nombre de points diacritiques infra ou supra-linéaires. Par exemple, le graphème [ح] peut donner lieu aux phonèmes ح /ħa :/, ج /ʤi:m/ ou خ /xa:ʔ/ en ajoutant ou en changeant le nombre ou la place des points.

[10] En arabe, les majuscules n'existent pas.

Selon leur point d'articulation fondamental, les consonnes sont des labiales, des dentales, des palatales, des vélaires ou des laryngales. Selon leur tenue, elles sont occlusives ([ب] /b/ ou [ك] /k/) ou spirantes ([ش] /ʃ/ ou [ز] /z/). Ces consonnes sont sourdes ([س] /s/ ; [ف] /f/) ou sonores ([ز] /z / ; [ت] /t/). Certaines sont compliquées d'une nasalisation ou d'une vélarisation.

Tableau 2 : Liste de l'alphabet arabe

Nom en arabe	Transcription phonologique	Isolée	Finale	Initiale	Médiale	Transcription phonologique
ألف	/ʔalif/	ء	ىٔ ، ـء ، ـؤ ، ـأ	أ	ـئـ ، ـؤ ، ـأ	/ʔ/
باء	/baːʔ/	ب	ـب	بـ	ـبـ	/b/
تاء	/taːʔ/	ة ، ت	ـة ، ـت	تـ	ـتـ	/t/
ثاء	/θaːʔ/	ث	ـث	ثـ	ـثـ	/θ/
جيم	/ǧiːm/	ج	ـج	جـ	ـجـ	/ǧ/
حاء	/ħaːʔ/	ح	ـح	حـ	ـحـ	/ħ/
خاء	/χaːʔ/	خ	ـخ	خـ	ـخـ	/χ/
دال	/daːl/	د	ـد	د	ـد	/d/
ذال	/ḏaːl/	ذ	ـذ	ذ	ـذ	/ḏ/
راء	/raːʔ/	ر	ـر	ر	ـر	/r/
زاي	/zaːj/	ز	ـز	ز	ـز	/z/
سين	/siːn/	س	ـس	سـ	ـسـ	/s/
شين	/ʃiːn/	ش	ـش	شـ	ـشـ	/ʃ/
صاد	/ṣaːd/	ص	ـص	صـ	ـصـ	/ṣ/
ضاد	/ḍaːd/	ض	ـض	ضـ	ـضـ	/ḍ/
طاء	/ṭaːʔ/	ط	ـط	طـ	ـطـ	/ṭ/
ظاء	/ḍaːʔ/	ظ	ـظ	ظـ	ـظـ	/ḍ/
عين	/ʔajn/	ع	ـع	عـ	ـعـ	/ʔ/
غين	/ʁajn/	غ	ـغ	غـ	ـغـ	/ʁ/
فاء	/faːʔ/	ف	ـف	فـ	ـفـ	/f/
قاف	/qaːf/	ق	ـق	قـ	ـقـ	/q/
كاف	/kaːf/	ك	ـك	كـ	ـكـ	/k/
لام	/laːm/	ل	ـل	لـ	ـلـ	/l/
ميم	/miːm/	م	ـم	مـ	ـمـ	/m/
نون	/nuːn/	ن	ـن	نـ	ـنـ	/n/
هاء	/haːʔ/	ه	ـه	هـ	ـهـ	/h/
واو	/waːw/	و	ـو	و	ـو	/w/
ياء	/jaːʔ/	ي	ـي	يـ	ـيـ	/j/

L'écriture arabe ne possède pas de lettres pour ses voyelles courtes. On trouve à la place de petites marques diacritiques placées au dessus ou au dessous des consonnes, soit le *fatha* : [ـَ], le ḍamma : [ـُ] et le kasra : [ـِ]. Il y a également le double fatha [ـً], le double damma [ـٌ] et le double kasra [ـٍ] connus sous le nom de *tanwin*. Ces trois signes sont utilisés et prononcés à la fin des noms et des adjectifs. Les deux diphtongues de l'arabe sont représentés par *ay* et *aw* (i.e., les mots *mawtun « mort »* et *sayrun « marche »*). En plus des signes diacritiques représentant les voyelles courtes, les voyelles longues et les diphtongues, il existe d'autres signes complémentaires :

(a) Sukoon [ْ] : ce signe indique l'absence de voyelles
(b) Shadda [ّ] : ce signe, écrit au dessus d'une lettre, marque la gémination ou le dédoublement de cette lettre.
(c) Madda [ٓ]: ce signe est écrit au dessus de la lettre alif pour indiquer son allongement.
(d) Hamza [ء] : ce signe note l'attaque ou la détente vocalique.

A l'exception de hamza , ces signes n'ont aucune valeur phonétique. Ils permettent seulement au lecteur de traiter certaines lettres dans des cas particuliers.

En arabe, la syllabe est soit ouverte, soit fermée (Blachère & Gaudefroy-Demombyne, 1975). La syllabe ouverte a pour structure : consonne + voyelle longue (e.g., [فِي][11] /fi:/) ou brève (e.g., ka/ta/ba/ [كَتَبَ][12]). La syllabe fermée se présente sous l'aspect : consonne + voyelle brève + consonne. C'est une syllabe longue (e.g., /kul/ « mange ! »). L'arabe classique ne connaît pas de syllabe fermée de type CCV[13]. Cependant, on peut rencontrer une telle syllabe en début de mot mais en reconstituant une syllabe initiale à l'aide d'un *alif* dit prosthétique servant de support à une attaque vocalique (e.g., [اِبْنٌ] /*ibnun/*[14]).

Au niveau orthographique, l'arabe est une langue alphabétique qui dispose de deux systèmes d'écriture, l'un vocalisé destiné surtout aux apprentis lecteurs et l'autre non vocalisé constituant l'écriture standard. Lorsque l'écriture est vocalisée, les voyelles sont représentées par des signes diacritiques surajoutés au squelette consonantique. La présence de la vocalisation crée des correspondances strictes entre graphèmes et phonèmes et procure ainsi une information phonologique et sémantique univoque.

[11] Dans
[12] Il a écrit
[13] Comme *station* en français
[14] Fils

Dans la forme non vocalisée, seul le squelette du mot composé de consonnes et de voyelles longues est transcrit, le mot est alors phonologiquement et sémantiquement ambigu. Par exemple, le mot non vocalisé «سلم» peut avoir différents sens : [سِلْمٌ] /*silmun*/[15], [سُلَّمٌ] /*sullamun*/[16] ou [سَلَّمَ] /*sallama*/[17]. Les lecteurs confirmés lisent les mots non vocalisés en s'appuyant sur le contexte. Par exemple, la phrase suivante contient un verbe qui ne peut pas être identifié hors contexte et sans voyelles puisqu'il est visuellement identique à plusieurs autres noms et verbes de différents sens : مر الرجل من هنا /*l'homme est passé par ici*/. Le premier mot de la phrase est un verbe. Sa prononciation correcte peut être obtenue par l'utilisation des voyelles. En l'absence de la vocalisation, le lecteur est obligé de recourir au contexte pour une lecture correcte. Il existe aussi en arabe des homographes hétérophoniques, c'est-à-dire des mots de différentes prononciations qui s'écrivent de la même façon en l'absence des voyelles. Par exemple, les mots [كُلّ] /*chaque*/ et [كُلْ] /*mange*/ ; [مرأة] /*femme*/ et [مرآة] /*miroir*/ sont identiques lorsqu'ils sont écrits sans voyelles. Ces homographes ne sont identifiés correctement que lorsqu'ils se présentent dans une phrase-contexte.

Pour l'écriture vocalisée, le lecteur débutant fait appel à une procédure séquentielle (Ammar, 2003). Dans un premier temps, les opérations portent sur le squelette consonantique pour lequel la prononciation est obtenue par la procédure d'assemblage. Dans un deuxième temps, en cas d'ambiguïté, les signes diacritiques sont utilisés. Cette stratégie, outre la connaissance des correspondances graphèmes-phonèmes, nécessite des compétences en analyse morphologique pour distinguer la racine des morphèmes vocaliques.

Au début de l'apprentissage, les enfants apprennent à lire dans une écriture totalement vocalisée. En Tunisie, l'écriture non vocalisée est introduite progressivement à l'école à partir de la cinquième année de scolarisation et la vocalisation est définitivement éliminée des manuels scolaires au niveau du cycle secondaire.

2.2. Système morphologique

La principale caractéristique du système morphologique arabe réside dans sa structure dérivationnelle. En langue arabe, le mot est constitué d'une racine purement consonantique qui donne l'idée générale. Les voyelles ne sont que des éléments de dérivation. Cette structure orthographique reflète le fonctionnement morphologique de la langue : le squelette

[15] Paix
[16] Escalier
[17] A dit bonjour

consonantique constitue la racine du mot, les voyelles constituent la flexion interne et fournissent une forme lexicale et grammaticale au mot. A partir de la racine, l'arabe construit par dérivation, au moyen des affixes et surtout de la flexion interne, l'ensemble de son vocabulaire. La racine arabe diffère donc nettement de la racine dans les langues indo-européennes où elle apparaît sous une forme syllabique, c'est-à-dire avec des voyelles dont les variations et les alternances jouent un rôle essentiel. Deux procédures de dérivation permettent de former un mot. Le mot peut être obtenu par concaténation d'un préfixe, d'un suffixe ou d'un infixe à la racine[18] ; il peut également être construit par superposition d'un morphème vocalique à la racine consonantique. Il s'agit d'une procédure de flexion interne qui permet de donner une forme lexicale ou grammaticale au mot. Les racines arabes sont dans leur immense majorité constituées par trois consonnes. Par exemple, les mots d'origine anglaise *bank* et *film* sont formés respectivement des racines b-n-k et f-l-m. Une même racine consonantique peut renvoyer ainsi à plusieurs représentations lexicales et phonologiques (voir tableau 3).

Un trait essentiel de la structure de l'arabe est que toute forme linguistique, dépourvue de ses éléments flexionnels, peut s'analyser fondamentalement en une racine et un schème (Pellat, 1960). Les schèmes sont générés par l'amalgame d'une racine fonctionnant comme unité lexicale consonantique virtuelle et d'un affixe ou plus (Baccouche, 2003). Un mot arabe n'est donc, abstraction faite des marques flexionnelles dont il est obligatoirement muni, que l'intersection en un point de la série des racines et de la série des schèmes[19]. Dans la plupart des cas, c'est le schème qui porte l'ensemble des déterminations du mot à l'exception de certains aspects de la fonction syntaxique. L'identification d'un schème d'une forme linguistique permet de reconnaître d'abord la partie du discours : nom ou verbe dont elle relève et, dans une grande mesure, la catégorie grammaticale à laquelle elle appartient (l'aspect du verbe, le genre et le nombre du nom, etc.). Mais elle relève dans un nombre important de cas, des déterminations d'ordre lexico-morphologique, par exemple s'il s'agit d'un nom abstrait ou concrêt ou d'un nom verbal d'action, d'un substantif, etc. Les racines ne subissent, de façon générale, aucune transformation du fait de leur insertion dans un schème. Les seules exceptions concernent les radicales W ou Y.

[18] Procédure analogue à celle des langues indo-européennes.
[19] Pour reconnaître un mot ou le chercher dans un dictionnaire, il faut connaître sa racine consonantique et son schème ; ce qui exige une connaissance minimale préalable du système linguistique arabe.

Tableau 3 : Exemple de dérivation à partir de la racine k-t-b

	Schèmes	Ecriture	Equivalent phonologique	Traduction
Racine k-t-b (idée d'écrire)	$R_1aR_2aR_3$-V	كَتَبَ	kataba	Il a écrit
	Ca-$R_1R_2uR_3$-V	يَكْتُبُ	jaktubu	Il écrit
	ma-$R_1R_2aR_3$	مَكْتَبٌ	maktab	Lieu où on écrit
	$R_1aaR_2iR_3$	كَاتِبٌ	ka:tib	Celui qui écrit
	ma-$R_1R_2uuR_3$	مَكْتُوب	maktu:b	Ce qui est écrit

R_1, R_2, R_3 : consonnes radicales.

C : consonne affixale variable.

V : voyelle affixale variable.

L'une des caractéristiques les plus importantes de l'arabe est qu'il est une langue à flexion dans laquelle des marques formelles, ajoutées au schème, constituent pour les verbes essentiellement des indices de personne et de mode et pour les noms des indices de fonction. Cette flexion confère aux mots une grande marge de liberté distributionnelle dans la phrase arabe[20]. Selon cette définition, l'arabe se classe comme une langue à morphologie extrêmement riche : le système dérivationnel se présente comme un jeu de construction basé sur le rôle sémantique de l'élément dérivé. L'association d'affixes à une racine donnée permet d'engendrer des mots avec des significations différentes, mais qui peuvent être classés comme mots de la famille du dérivant. Par exemple, du verbe *kasara* signifiant « casser » sont dérivés, par doublement de la consonne /s/, le verbe *kassara* « casser en mille morceaux », et par ajout du préfixe /in/, le verbe *inkasara* « se casser ». La définitude est réalisée, elle aussi, par un morphème, préfixé (al-) ou suffixé (-n), et non par des mots autonomes. Dans le script, la marque du défini est réalisée par des lettres tandis que la marque de l'indéfini est réalisée par un signe diacritique fusionné au signe de la voyelle courte, c'est le signe appelé tanwin :

Exemple : ***al****+awlaad +u* ***vs*** *ʔawlaad+u+****n***
DEF + (N) + NOM (N) + NOM + **INDEF**
Les enfants **Des** enfants

L'arabe a un système très régulier de conjugaison des verbes et de modification de leurs racines pour indiquer les variations du sens de base. Les noms et les adjectifs sont moins

[20] En typologie des langues, on appelle langue flexionnelle une langue dans laquelle les mots changent de forme (soit le nombre des noms, soit le temps verbal) selon les rapports grammaticaux qu'ils entretiennent avec les autres mots. Certains mots modifient donc leur forme (sonore et/ou visuelle) ; on dit d'eux qu'ils subissent le jeu de la flexion.

réguliers sur le plan de la formation et ils présentent plusieurs structures différentes de formation du pluriel. Les pluriels internes s'obtiennent en modifiant la forme du nom singulier, par exemple, les pluriels des mots empruntés *bank* [بنك] et *film* [فلم] sont respectivement *bunu:k* [بنوك] et *afla:m* [أفلام].

3. La situation de diglossie

La diglossie est, selon Ferguson (1959), une situation où sont en usage deux langues apparentées génétiquement et structurellement et dont les distributions fonctionnelles sont complémentaires. Contrairement au bilinguisme, la diglossie se caractérise par une situation relativement stable dans laquelle les deux langues, bien qu'utilisées de façon complémentaire, ne bénéficient pas du même statut social, l'une ayant un statut socioculturel supérieur à l'autre. Dans les pays arabes, la situation de diglossie traduit la coexistence de deux variétés ou formes linguistiques parentes (Ben Rejeb, 2001). La première variété, l'arabe dialectal est la langue de la communication quotidienne. La seconde variété, l'arabe standard, est la langue des enseignements scolaires, utilisée dans le cadre officiel. Si le dialectal domine à l'oral, le standard s'impose plus à l'écrit, car la transcription du dialectal n'est pas normalisée (Baccouche, 2003).

En Tunisie, Baccouche (1994) distingue deux types d'arabe : l'arabe littéral et l'arabe parlé. Il distingue deux niveaux de l'arabe littéral : L'arabe littéral classique utilisé surtout dans la poésie traditionnelle, les romans de haute tenue stylistique et le discours religieux et l'arabe littéral moderne dont le prototype est la langue journalistique influencée par le français au niveau lexical et syntaxique. Pour l'arabe dialectal, il distingue également deux niveaux : un niveau populaire et familier qui exprime les besoins quotidiens et un niveau intellectualisé, utilisé surtout dans la conversation des lettrés et dans la plupart des émissions de radio et de télévision. Baccouche (1994) parle également de niveaux intermédiaires et propose de distinguer *l'arabe médian*, niveau intermédiaire entre le littéral et le dialectal avec des interférences et des emprunts réciproques et le *francarabe* qui représente un mélange d'arabe dialectal tunisien et de français.

Cette situation sociolinguistique particulière rend l'apprentissage scolaire de la langue arabe plus contraignant et plus laborieux pour les apprentis lecteurs. En effet, l'enfant tunisien doit, en entrant à l'école, effectuer un double passage : de l'oral à l'écrit et du registre dialectal au registre standard. Des différences importantes existent entre les deux registres sur les plans

morphologique, sémantique, syntaxique et surtout phonologique. En effet, l'arabe dialectal diffère de l'arabe standard par une syntaxe simplifiée, un lexique plus riche en vocables étrangers et une phonologie altérée. Par exemple, certaines voyelles comme /e/ et /o/ existent en arabe dialectal mais pas en arabe standard. En arabe dialectal, certains mots commencent par deux consonnes consécutives ou par une consonne et un sukoon, ce qui est illégal en arabe standard (voir tableau 4). Les deux formes utilisent différentes flexions (comme les marques du pluriel) et différentes contraintes dans l'ordre du mot dans phrase. Il existe également des mots en arabe standard qui n'ont pas la même représentation phonologique en arabe dialectal (le mot /*matbaxun/* [مطبخ] qui désigne une cuisine en arabe standard devient /*ku:ʤina*/ [كوجينة] en arabe dialectal).

Si dans d'autres contextes, la lecture contribue au renforcement et à l'enrichissement d'un lexique préexistant, le lecteur tunisien débutant découvre, en même temps, la forme écrite du mot et le son qui lui correspond. La langue écrite, pour lui, n'est ni un prolongement, ni l'aspect écrit de la langue parlée. En effet, c'est sur ses connaissances linguistiques provenant de son expérience de l'oral que l'enfant va s'appuyer pour identifier les mots écrits. Dés lors, la capacité métaphonologique primordiale dans l'accès à l'écrit va porter sur la langue orale, l'arabe dialectal. Le changement de répertoire, du dialectal au standard, constituerait ainsi une source potentielle de difficultés pour l'apprentissage de la lecture qui se traduirait par des interférences phonologiques entre les deux registres (Ammar, 1997 ; Belajouza & Snoussi, 2006 ; Snoussi, 2007).

Tableau 4 : Exemples de diglossie

Mots en français	**Prononciation en arabe standard**	**Prononciation en arabe dialectal**
Livre	/kita:b/	/kta:b/
château	/qaşr/	/qşar/
Cheval	/ħişa:n/	/ħşa:n/

Plusieurs recherches ont tenté d'étudier l'impact de la situation de diglossie sur les processus d'identification des mots mis en jeu par le lecteur arabophone au cours de l'apprentissage de la lecture. Certaines recherches ont attribué la grande difficulté dans l'apprentissage de la lecture en arabe à la situation de diglossie (Feitelson, 1993 ; Saiegh-Haddad, 2003 ; 2004),

d'autres études montrent en revanche que la diglossie peut jouer un rôle facilitateur dans le développement de la conscience phonologique et de la lecture de mots isolés (Ibrahim, Eviatar & Aharon-Peretz, 2007).

Ammar (1997) a observé que les performances des bons lecteurs débutants en arabe dans les tâches de segmentation phonémique sont nettement inférieures à celles observées sur les lecteurs débutants en anglais ou en hébreu. Selon l'auteur, ces difficultés sont imputables à la contrainte que constitue la diglossie. Les changements phonologiques auxquels les enfants sont confrontés lors du passage du répertoire dialectal au répertoire standard ont gêné les capacités d'identification des phonèmes. Par ailleurs, l'étude de Feitelson & al. (1993) sur des enfants palestiniens préscolarisés en jardin d'enfants montre que le problème du passage de l'arabe dialectal à l'arabe standard peut être réduit par une éducation préscolaire appropriée susceptible de familiariser les enfants avec la phonologie de l'arabe standard. Les auteurs ont soumis les enfants à des séances d'apprentissage de comptines en arabe standard. L'évaluation a montré que les performances de ce groupe s'avère meilleures que celles d'un groupe contrôle. Les enfants ont acquis un vocabulaire riche en arabe standard et syntaxiquement plus correcte. Le développement d'un lexique phonologique riche en arabe standard permettrait ainsi de minimiser les effets d'interférence.

Abu-Rabia (2000) a comparé les performances en compréhension écrite chez des élèves de 1ère et 2ème années qui ont été auparavant sensibilisés à l'arabe standard pendant les années préscolaires, à un groupe contrôle. Les résultats montrent que l'exposition précoce des enfants arabes préscolaires à l'arabe standard améliore leurs performances en compréhension écrite deux années plus tard. Les élèves qui ont été exposés à l'arabe standard font preuve d'une meilleure compréhension écrite que les élèves qui n'ont été en contact qu'avec l'arabe parlé.

Saiegh-Haddad (2003) a examiné l'effet des différences phonologiques entre l'arabe dialectal et l'arabe standard sur l'acquisition des compétences phonémiques et des compétences en identification de mots chez 23 enfants palestiniens de jardins d'enfants et 42 enfants de 1ère année. Trois questions ont été posées dans cette étude :

- Est- ce que la distance linguistique entre l'arabe dialectal et l'arabe standard interfère avec l'acquisition des processus de base de la lecture en arabe standard ?
- Quelle structure diglossique (phonologique, syntaxique, morphosyntaxique ou lexicale) interfère avec l'acquisition de la lecture ?

- Parmi les compétences liées à l'apprentissage de la lecture (conscience phonémique, reconnaissance de mots, rapidité de la lecture ou compréhension écrite), lesquelles sont sensibles à la situation de diglossie ?

L'épreuve de conscience phonémique comporte deux tâches : isolation du phonème initial (40 items) et du phonème final (40 items). Le niveau de lecture est évalué par une tâche de lecture de pseudo-mots (40 items). Les résultats montrent que les enfants préscolaires et ceux de 1ère année trouvent des difficultés à isoler les phonèmes en arabe standard. Cette difficulté est encore plus importante si ce phonème est présent en position initiale dans une structure syllabique propre à l'arabe standard. Les enfants arabes ont des difficultés dans la manipulation du phonème consonantique pré-vocalique que le phonème consonantique postvocalique final. Mais d'une façon générale, les jeunes palestiniens de 1ère année réussissent mieux ces tâches de suppression phonémique que les élèves du préscolaire. Les élèves de 1ère année semblent profiter de leur exposition à l'arabe standard et surtout de l'instruction explicite de la lecture. Cette étude montre que la situation de diglossie et plus particulièrement la distance phonologique entre les deux variétés de l'arabe a un effet sur l'habileté de décodage chez les apprentis lecteurs arabophones.

Dans une autre étude, Saiegh-Haddad (2004) a essayé d'étudier l'effet de la distance phonémique et lexicale entre l'arabe dialectal et l'arabe standard en utilisant cette fois différents types de mots : des mots en arabe standard (AS) et en arabe dialectal (AD) et des pseudo-mots (PM). Les résultats montrent que les phonèmes en AS, qui sont absents de l'expérience langagière des enfants, sont plus difficiles à isoler que les phonèmes en AD aussi bien chez les enfants préscolaires que ceux de 1ère année. Les données montrent également que les enfants préscolaires trouvent plus de difficulté à manipuler les mots et les phonèmes en AS que les PM. Cependant, les enfants de 1ère année, qui auraient bénéficiés d'un an d'exposition à l'AS, ont développé des représentations phonologiques des mots et des phonèmes dans ce registre, ce qui explique leurs meilleures performances dans la manipulation des phonèmes en AS. Ainsi, les jeunes arabophones doivent simultanément maîtriser des représentations phonologiques de certaines structures qui ne sont pas disponibles dans leur langage oral et découvrir comment ces structures peuvent être intégrées dans ce système orthographique.

Ibrahim, Eviatar & Aharon-Peretz (2007) ont testé l'effet positif du bilinguisme précoce, qui affecte la conscience phonologique, sur les performances en lecture. Ils ont comparé trois

groupes de première année : un groupe de 20 monolingues hébreux, un groupe de 19 bilingues russe-hébreu et un groupe de 20 enfants arabophones dans des tâches d'évaluation des connaissances phonologiques, du niveau de vocabulaire et de lecture. La conscience phonologique a été évaluée par une tâche de détection du phonème initial, une tâche de détection du phonème final, et une tâche de suppression syllabique et phonémique en position initiale, médiane et finale. Pour l'évaluation du niveau de vocabulaire, les auteurs ont utilisé le test de définition de mots du WISC-R. le niveau de lecture a été évalué par deux épreuves : lecture de texte et lecture de mots et de pseudo-mots isolés. Les résultats indiquent que l'expérience linguistique de chaque groupe influence sa conscience phonologique. Les bilingues ainsi que les arabophones manipulent mieux les unités linguistiques que les monolingues hébreux. Concernant la tâche de vocabulaire, les scores se sont inversés au profit des monolingues hébreux, alors que les performances des deux autres groupes ne se différencient pas de façon significative. Les arabophones, considérés comme des bilingues, montrent également un léger avantage dans la lecture de mots et de pseudo-mots par rapport aux deux autres groupes lisant en hébreu. Mais l'analyse des corrélations dégage un faible coefficient entre les connaissances phonologiques et les performances en lecture. L'analyse de régression montre que la conscience phonologique explique 60 % de la variance des performances en lecture chez les monolingues et les bilingues hébreux alors qu'elle explique seulement 30% de la variance chez les lecteurs arabophones. Les auteurs expliquent ces différences de performances, malgré des scores similaires en phonologie, par la complexité de l'orthographe arabe qui exigerait plus de ressources cognitives que l'hébreu.

Les travaux de Ammar (1997) ont mis en évidence l'existence de deux registres lexicaux, en arabe dialectal et en arabe standard, qui s'organisent différemment en mémoire. Le premier opère dans le traitement phonologique et sémantique du langage oral et le second est réservé au traitement du langage écrit. Cependant, au début de l'apprentissage, la mise en œuvre du registre lexical réservé à l'arabe dialectal au cours de l'activité de lecture donne lieu à des interférences phonologiques avec la langue de l'écrit[21]. Dans le même sens, la recherche de Derguini & Mekideche (2003) menée en Algérie montre que l'apprenti lecteur arabophone dispose, à partir du troisième mois de scolarité, de deux lexiques mentaux. Le premier en arabe dialectal, déjà élaboré et stocké avant l'entrée à l'école et le second en arabe standard, en cours d'élaboration via l'apprentissage scolaire, et qui est considéré par l'institution scolaire comme le pré-requis à l'acquisition du langage écrit. A partir de la 4ème année, le sujet

[21] Les élèves auront tendance à donner des réponses combinant des éléments phonologiques des deux registres.

censé être un lecteur expert, disposerait d'une possibilité équivalente de recourir à l'un et à l'autre lexique.

4. Spécificités de l'apprentissage de la lecture en langue arabe

Les recherches sur l'évolution de la compétence lexique en anglais (*cf.* chapitre 1) décrivent trois étapes ou stratégies successives (logographique, alphabétique et orthographique) par lesquelles passent nécessairement les enfants pour parvenir à la maîtrise de la lecture dans une écriture alphabétique. Ces recherches ont confirmé le rôle central de la médiation phonologique et la liaison étroite entre les habiletés métaphonologiques et le niveau de compétence en identification des mots. Dans cette perspective, nous pouvons nous poser les questions suivantes :

Est-ce que l'évolution des stratégies d'identification de mots en arabe est similaire à celle observée en anglais ?

Quel est l'impact des contraintes imposées par la structure morphologique et le double système d'écriture de la langue arabe standard sur l'apprentissage de la lecture ?

4.1. L'évolution des stratégies de lecture en arabe

Les premiers travaux sur l'évolution des stratégies d'identification des mots écrits en arabe (Ammar, 1997 ; 2002 ; 2003) ont pu montrer que la reconnaissance des mots en arabe est fondée sur le morphème consonantique. Les opérations de codage grapho-phonologique concernent en premier les graphèmes de la racine du mot (les consonnes) et les informations vocaliques ne sont utilisées que s'il est nécessaire de résoudre une ambiguïté phonologique ou sémantique.

Dans le même sens, Taouk & Coltheart (2004) ont étudié les stratégies d'acquisition de la lecture en arabe chez trois groupes d'enfants âgés de 8, 9 et 11 ans et chez des adultes âgés de 22 à 43 ans. Les participants devaient lire quatre listes de non-mots en arabe standard. La liste 1 représente des non-mots vocalisés. La liste 2 représente des non-mots vocalisés dont la perturbation se situe au niveau des voyelles. La liste 3 représente des non-mots non vocalisés constitués par trois consonnes, deux d'entre elles sont écrites en double. La liste 4 représente des mots rares non vocalisés. Les résultats indiquent que les enfants sont capables d'utiliser la

stratégie alphabétique pour lire les différents types de non-mots arabes avec une supériorité pour les enfants les plus âgés et de niveau supérieur. Les enfants obtiennent des performances meilleures dans la lecture de la liste 1 par rapport à la liste 2. Ces données laissent supposer que les enfants les plus jeunes s'appuient sur les indices visuels saillants (les consonnes) pour accéder au mot dans leur lexique mental. L'observation des résultats obtenus à la liste 4 et la liste 5 qui présentent des mots non vocalisés, montre une utilisation plus fréquente du « sukoon »[22] à la fin du mot en fonction de l'âge et du niveau scolaire. En effet, l'utilisation du « sukoon » augmente systématiquement et de façon significative du grade 3 jusqu'à l'âge adulte alors que la fréquence d'utilisation de la voyelle « fatha » diminue systématiquement du grade 3 jusqu'à l'âge adulte. L'utilisation plus fréquente du « sukoon » suggère que les lecteurs confirmés tendent à ignorer les voyelles à la fin des mots en les remplaçant par le « sukoon » alors que les lecteurs débutants traitent toutes les voyelles dans un mot. En considérant que le rôle principal des voyelles finales en arabe est d'indiquer la fonction grammaticale et le sens du mot, ces données montrent que les lecteurs s'appuient plus sur le contexte et moins sur la voyelle finale avec l'avancement dans la scolarité. Ceci est en contradiction avec les résultats obtenus sur l'anglais et le français qui ont montré que les lecteurs débutants et les faibles lecteurs s'appuient plus que les lecteurs experts sur le contexte pour identifier les mots (Nicholson, 1991 ; Perfetti, Goldman & Hogoboam, 1979 ; Sprenger-Charolles & Khomsi, 1989 ; Stanovich, 1981).

Les auteurs finissent par conclure que les trois phases qui caractérisent l'évolution des stratégies de la lecture en anglais existent bien en arabe. Au début de l'apprentissage, les enfants utilisent la stratégie logographique basée sur les indices graphiques saillants, ils s'appuient sur les consonnes pour identifier les mots et négligent les voyelles. Les lecteurs en grade 6 utilisent la stratégie alphabétique dans la lecture des non-mots vocalisés. L'utilisation systématique de la voyelle « fatha » à la fin des mots correspond à une étape précoce de recodage phonologique ou l'identification du mot s'appuie sur la correspondance lettre-son. Avec l'avancement dans la scolarité, les lecteurs experts en arabe utilisent la stratégie orthographique. Leur performance est inférieure lorsque les mots présentent une perturbation orthographique au niveau des consonnes (liste 3). L'utilisation systématique du « sukoon » correspond au début de la phase orthographique ou l'identification du mot est directement reliée au contexte. Ainsi, malgré les différences entre les deux orthographes, l'identification du mot en arabe, comme en anglais, est caractérisée par l'existence d'une phase initiale

[22] Ce signe indique l'absence de voyelles.

logographique suivie par une phase phonologique. Ensuite, il y a une transition graduelle vers une phase orthographique.

Taouk & Coltheart (2004) considèrent que les caractéristiques du système d'écriture arabe facilitent l'utilisation de la stratégie alphabétique. En effet, certaines difficultés associées à l'utilisation du recodage grapho-phonologique chez l'apprenti lecteur anglais n'existe pas en arabe telles que l'incapacité de distinguer entre homophones (exemple « *sail* » et « *sale* ») et de prononcer correctement les mots irréguliers et les homographes hétérophréniques. D'un autre côté, certaines caractéristiques de la langue arabe posent un grand problème pour les apprentis lecteurs arabophones en comparaison avec les apprentis lecteurs anglais : la situation de diglossie, la prééminence des consonnes par rapport aux voyelles, la grande ressemblance entre certaines lettres et la variation de la forme des lettres selon leur position dans le mot.

Ibrahim, Eviatar & Peretz (2002) ont montré que la grande complexité de l'orthographe arabe ralentit les processus d'identification de mots. L'étude a porté sur 30 sujets bilingues dont l'arabe est la langue maternelle et l'hébreu est la langue seconde. Les résultats révèlent que les lettres sont identifiées plus rapidement en hébreu qu'en arabe. Les auteurs expliquent cette difficulté dans l'identification des lettres en arabe par rapport à l'hébreu par la grande complexité de l'orthographe arabe qui augmente la charge cognitive et ralentit les processus d'identification des mots. En effet, ils ont trouvé que les élèves de première année font plus d'erreurs sur les affixes (voyelles et consonnes), qui dépendent des règles syntaxiques et de la fonction des mots dans la phrase lors de la lecture de textes, que sur les mots eux-mêmes.

Azzam (cité par Taouk & Coltheart, 2004) a examiné les erreurs de lecture chez 150 enfants âgés de 6 à 11 ans apprenant à lire en arabe. Les résultats montrent que les erreurs de lecture concernent principalement les voyelles qui fournissent l'information phonologique. Au début de l'apprentissage de la lecture, les enfants passent par une phase logographique caractérisée par une identification de mots basée sur des indices graphiques saillants. Avec l'avancement dans la scolarité, les observations révèlent une utilisation massive des règles de CGP dans tous les grades, ce qui est du à la transparence des CGP en arabe. Les résultats montrent également que le nombre d'erreurs diminue en fonction du niveau scolaire, ce qui indique que les enfants ont recours à la stratégie orthographique. Ainsi, les enfants arabes, comme leurs homologues anglais, semblent passer d'une phase alphabétique basée sur les CGP à une phase

orthographique basée sur une reconnaissance directe au fur et à mesure qu'ils avancent dans la scolarité.

A l'aide d'une épreuve d'identification de mots en contexte imagé, Ammar (1997) a trouvé que les perturbations sur les consonnes étaient beaucoup mieux détectées que les perturbations sur les voyelles aussi bien chez les bons lecteurs que chez les faibles lecteurs. Il conclut que l'identification des mots en arabe est fondée sur le squelette consonantique. Cette procédure implique une analyse morphologique préalable, aboutissant à une décomposition du mot en racine consonantique et morphèmes vocaliques. Il semble que les opérations de codage grapho-phonologique concernent d'abord la racine consonantique et que les informations vocaliques sont analysées dans un deuxième temps et seulement s'il est nécessaire de résoudre une ambiguïté sémantique ou phonologique. Selon Ammar, ces opérations se déroulent en deux cycles conformément au modèle de Berent & Perfetti (1995). Durant le premier cycle, seules les consonnes sont analysées. Le décodage de la racine permet l'identification des mots familiers. Le second cycle qui porte sur l'exploitation des voyelles permet l'identification des mots rares ou des pseudo-mots.

4.2. Rôle des voyelles et du contexte

Le rôle des voyelles et du contexte dans l'apprentissage de la lecture en arabe a été abondamment étudié au cours des expérimentations réalisées par Abu Rabia (2001, 1998, 1997, 1996 ; Abu Rabia & Siegel, 1995). Dans une étude avec des bons lecteurs et des faibles lecteurs en arabe, Abu Rabia (1997) a tenté de tester l'effet des voyelles et du contexte sur les performances en lecture. Les sujets devaient lire 15 paragraphes, 60 phrases et 210 mots en arabe dans trois conditions différentes : texte totalement vocalisé, texte partiellement vocalisé et texte non vocalisé. Les résultats obtenus dans la lecture des paragraphes et des phrases indiquent que la présence des voyelles facilite l'identification des mots aussi bien chez les bons lecteurs que les faibles lecteurs. Ces données sont en contradiction avec celles observées en hébreu (Frost & Bentin, 1992 ; Koriat, 1984) où la présence des voyelles ne semble pas accélérer la décision lexicale et avec celles observées par Ammar (1997) en arabe. Cependant, l'auteur n'observe pas de différences significatives entre la lecture des textes vocalisés et partiellement vocalisés. Ceci peut être expliqué par le recours des lecteurs au contexte pour compenser l'absence de certaines voyelles (Abu Rabia & Siegel, 1995). Les résultats obtenus dans la lecture des mots isolés montrent que les bons et les faibles lecteurs sont plus performants dans la lecture des mots vocalisés, ce qui confirme l'importance des voyelles et

du contexte dans l'apprentissage de lecture en arabe. Ces résultats divergent des données recueillies sur l'anglais (Stanovich, 1979) et sur le français (Sprenger-Charolles & Khomsi, 1989) qui ont montré que les bons lecteurs utilisent le contexte à un moindre degré que les faibles lecteurs.

Dans une autre expérience, Abu Rabia (1998) a étudié l'effet de la vocalisation et du contexte en utilisant différents types de matériel : poèmes, textes coraniques, articles de journaux et des contes. Les sujets avaient pour tâche de lire des listes de mots et de paragraphes en version vocalisée et non vocalisée. Un effet significatif de la vocalisation et du contexte a été observé aussi bien chez les bons lecteurs que les faibles lecteurs et pour chaque type de texte.

La lecture des mots isolés non vocalisés est problématique en arabe aussi bien pour les bons et les faibles lecteurs. La seule alternative pour la lecture correcte de ces mots consiste à les présenter dans une phrase ou un paragraphe. Selon Abu Rabia (1997), « *arabic is perhaps the only language in the world in which readers must first understand the sentence to recognize the word* » (p. 480). L'apprentissage de la lecture doit être considéré comme un processus interactif et dynamique impliquant l'identification des mots et le contexte (*cf.* figure 3). Si le texte est non vocalisé, la seule issue pour assurer une identification de mots correcte est de s'appuyer sur le contexte. Ces résultats convergent avec ceux trouvés par Ammar (1997) qui montrent que les faibles lecteurs ne compensent pas leur déficit lexique par un plus grand recours au contexte. En effet, ils obtiennent de basses performances dans la détection des perturbations sémantico-contextuelles et s'avèrent moins performants dans la détection des perturbations orthographiques portant sur les signes vocaliques que les bons lecteurs. Les enfants faibles lecteurs se trouvent ainsi dans une situation de surcharge cognitive qui ne leur permet pas de traiter à la fois l'information orthographique émanant du squelette consonantique, l'information contextuelle, et encore moins l'information procurée par les signes diacritiques.

Dans une expérience plus récente, Ammar (2003) confirme le traitement séquentiel des consonnes et des voyelles au cours de l'identification des mots écrits en arabe. Il a étudié les prononciations produites par les élèves de 2ème, 4ème et 6ème année dans la lecture à haute voix de pseudo-mots vocalisés. Les résultats permettent de constater que l'information phonologique procurée par les signes diacritiques n'est pas ignorée dans la lecture à voix haute. Elle est utilisée pour la prononciation des pseudo-mots dont la perturbation porte sur les consonnes ou sur les voyelles. En revanche, Ammar observe que plus de 40% des enfants

se trompent dans la prononciation des voyelles. En revanche, la prononciation des consonnes est le plus souvent correcte.

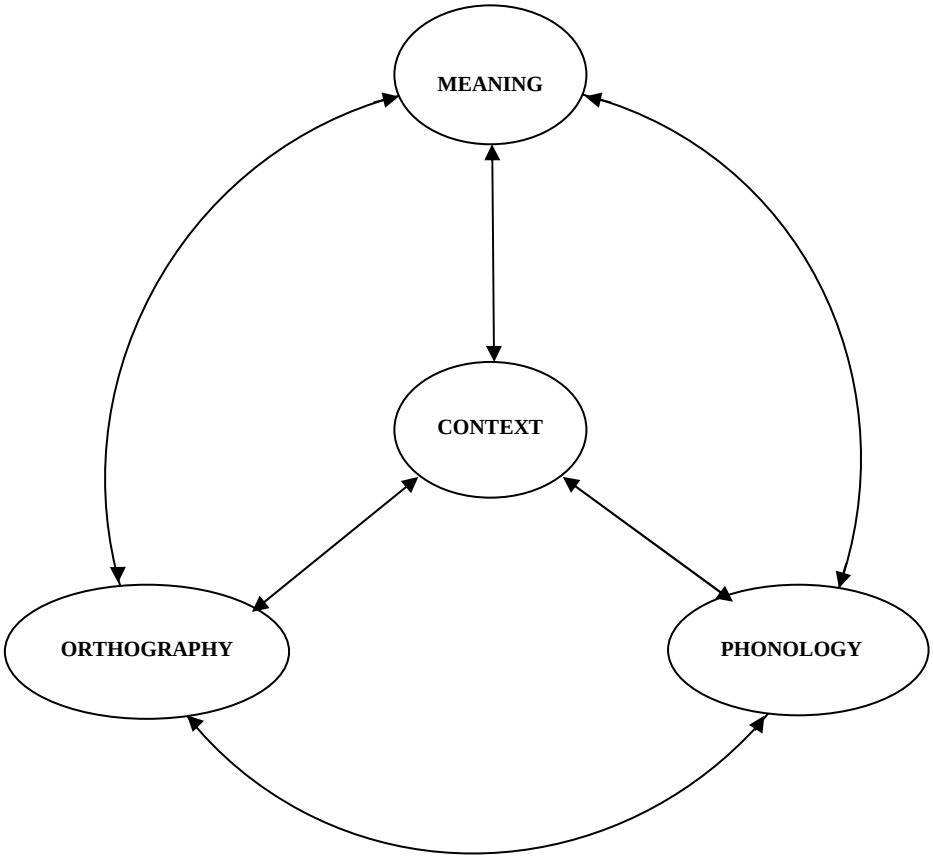


Figure 3 : Modèle de lecture chez le lecteur habile en arabe (d'après Abu-Rabia, 1997)

L'étude de Roman, Pavard & Asselah (1985) avait pour objectif d'étudier et d'analyser les stratégies utilisées par les lecteurs arabes pour intégrer les mots de phrases ambigües. En arabe non vocalisé, l'absence de voyelles prive le lecteur des informations spécifiant le cas du verbe (actif ou passif). L'ambigüité des phrases utilisées dans cette expérience porte sur le verbe, premier élément de la phrase canonique en arabe. En version non vocalisée, le verbe ne porte pas de marque casuelle : le lecteur ne peut donc pas définir sans ambigüité la flexion casuelle du verbe (actif ou passif) et la fonction du syntagme nominal qui suit le verbe. Dans une première expérience, les auteurs ont utilisé des phrases non vocalisées dont le verbe est ambigu. La représentation du verbe correspondra ou non, selon les conditions, à la représentation la plus fréquente selon l'usage courant. Dans une seconde expérience, ils ont utilisé les mouvements oculaires pour identifier les stratégies adoptées par le lecteur arabe pour accéder aux mots et résoudre les ambigüités locales. Les deux expériences montrent que

le lecteur code le verbe ambigu soit sur la base de la forme casuelle la plus courante dans la langue (forme active), soit sur la base de la forme casuelle la plus courante du verbe considéré. En arabe non vocalisé, le lecteur privilégie, en lecture orale, la forme active (80% des phrases passives sont lues à l'actif et 46% des phrases actives sont lues au passif). Lorsque les phrases sont vocalisées, le codage phonologique est correct dans 95% des cas alors que lorsqu'elles ne sont pas vocalisées, les sujets ne choisissent la bonne interprétation vocalique que pour 60% des phrases. Pour les phrases vocalisées, le passif est aussi bien identifié que l'actif. La vocalisation est donc utilisée pour effectuer l'accès au lexique de la racine ambigüe pendant la lecture. L'analyse des mouvements oculaires met en évidence que le temps alloué au traitement du verbe passif vocalisé est en moyenne de 600 ms contre 300 ms pour le verbe non vocalisé. Les régressions s'effectuent lorsque les phrases sont ambigües (non vocalisées) et sont plus importants sur les derniers éléments de la phrase que sur le verbe. Ce résultat semble signifier que les sujets préfèrent réanalyser la proposition désambigüisante que réinterpréter la structure syntaxique de la phrase.

L'ensemble de ces études a mis en évidence que les mécanismes d'acquisition de la lecture en arabe sont affectés par la structure particulière de la langue.

5. Conclusion

Comme on a pu le constater précédemment, les données sur l'apprentissage de la lecture ont été développées à partir de la langue anglaise avec une visée de généralisation à toutes les langues. Or, le système orthographique de l'arabe se caractérise par de fortes spécificités. Une des principales caractéristiques de l'arabe est la structure bidimensionnelle du mot. Contrairement à l'anglais où les voyelles font partie de l'alphabet et sont représentées par des lettres spécifiques, en arabe les voyelles sont représentées par des signes diacritiques. La prééminence des consonnes est renforcée par le système d'écriture. Ainsi, la langue arabe repose sur un système qui peut être vocalisé ou non. L'écriture vocalisée, utilisée au début de l'apprentissage de la lecture, est phonologiquement transparente. L'écriture non vocalisée se caractérise en revanche par son opacité liée à l'absence de signes diacritiques. Certains chercheurs ont postulé que le lecteur a recours au contexte pour résoudre l'ambigüité de certains mots (Abu-rabia, 1997 ; 1998 ; 2001 ; Abu-Rabia & siegel, 1995). Les caractéristiques orthographiques et morphologiques de l'arabe font que l'analyse morphologique est nécessaire lors de la reconnaissance des mots écrits, et qu'elle joue un rôle plus immédiat que dans les langues latines (Gombert, 2003).

Une autre spécificité de la langue arabe provient du fait que la forme des lettres varie selon leur position dans le mot. La majorité des lettres arabes ont trois formes différentes selon qu'elles sont isolées, au début, au milieu ou à la fin des mots. De plus, différentes lettres sont similaires et la distinction entre elles est difficile. L'identification de ces différentes lettres et la connaissance de leur prononciation en différentes positions demande une charge cognitive considérable et constitue une difficulté pour l'identification de mots (Abu-Rabia, 2001). Par ailleurs, la situation de diglossie constitue une caractéristique fondamentale du contexte d'apprentissage dans les pays arabes. Sur le plan oral, la diglossie représente une contrainte phonologique supplémentaire aux apprentis lecteurs (Saiegh-Haddad, 2004 ; 2003). Ainsi, les spécificités linguistiques des langues sémitiques sont susceptibles de renouveler de façon importante les modèles d'acquisition de la lecture qui étaient jusqu'alors étayés par les données obtenues dans les langues à écriture latine. Dans le chapitre suivant, nous présenterons les facteurs qui sont susceptibles d'expliquer la réussite ou l'échec dans l'apprentissage de la lecture. Certains d'entre eux constituent des prédicteurs puissants des performances ultérieures en lecture.

Chapitre 3 : Les déterminants de la réussite dans l'apprentissage de la lecture

L'apprentissage de la lecture est un objet d'étude d'une grande complexité du fait de ses dimensions multiples. Il fait ainsi intervenir aussi bien des habiletés spécifiques au domaine particulier du traitement de l'information écrite (connaissances phonologiques, connaissances morphologiques, connaissance des lettres, etc.) que des habiletés cognitives beaucoup plus générales (mémoire, intelligence, attention, etc.) qui interviennent dans bien d'autres activités d'apprentissage.

L'acquisition de la langue écrite est ainsi dépendante de contraintes individuelles ou internes liées aux caractéristiques cognitives du sujet apprenant et de contraintes externes (Khomsi, 1995). Au niveau du sujet apprenant, l'acquisition de la lecture met en jeu des compétences cognitives générales (capacités intellectuelles, mnésiques, attentionnelles et d'automatisation) ainsi que des habiletés métalinguistiques, en particulier les compétences métaphonologiques. Quant aux contraintes externes, elles concernent les choix pédagogiques relatifs à l'enseignement de la langue écrite et les spécificités de la langue d'apprentissage.

1. Les connaissances phonologiques

L'apprentissage de la lecture dans un système alphabétique est un processus complexe qui requiert et implique de multiples habiletés linguistiques et non linguistiques, notamment l'habileté à réfléchir sur le langage. Cette capacité métalinguistique désigne l'activité de réflexion sur le langage et le contrôle délibéré opéré par le sujet sur les processus en œuvre dans le traitement du langage (Gombert, 1990). De nombreuses études ont été consacrées aux relations entre l'apprentissage de la lecture et les habiletés phonologiques, c'est-à-dire la capacité à opérer une analyse phonologique du langage oral qui conduit l'enfant à repérer puis isoler mentalement les éléments de la parole constituant un mot parlé (Demont, 2003 ; Demont & Gombert, 2007 ; Ecalle & Magnan, 2007a ; 2007b ; 2003 ; 2002 ; Gombert, 1990, 1992a ; Tunmer, 1989 ; Mann, 1989). Les habiletés phonologiques ont été étudiées dans différentes tâches consistant à comparer ou manipuler les sons indépendamment du sens. Parmi les tâches les plus fréquemment utilisées, on peut citer : la tâche d'identification de phonèmes (quel est le premier / le dernier son dans le mot « chat » ?) ; la tâche de segmentation de phonème (quels sons entendez-vous dans le mot « chat » ?) ; la tâche de

comptage de phonèmes (combien de sons entendez-vous dans le mot « chat » ?) et la tâche d'identification du phonème manquant (quel son entendez-vous dans le mot « char » qui manque dans le mot « chat » ?) (Gombert & Colé, 2000).

On peut donc distinguer différents niveaux d'abstraction dans les connaissances phonologiques correspondant à différents degrés de compétence mettant en jeu le type de traitement, la nature et la taille des unités traitées (Gombert & Colé, 2000 ; Ecalle, Magnan & Bouchafa, 2002). Les compétences phonologiques ont été étudiées, d'une part, chez des enfants prélecteurs dans le but de mieux comprendre comment avant l'apprentissage institutionnel de la lecture, le jeune enfant se représente les sons de la parole et d'autre part, chez les lecteurs débutants afin de mieux comprendre la façon dont les enfants vont acquérir cette conscience phonologique.
Les travaux présentés dans cette partie apportent quelques précisions sur le rôle des connaissances phonologiques lors de l'apprentissage de la lecture. Après avoir précisé comment se développe la conscience phonologique, nous présenterons des travaux sur les liens classiquement étudiées entre celle-ci et l'apprentissage de la lecture.

1.1. Les unités de traitement phonologique

Les processus de recodage phonologique sont considérés comme un puissant outil utilisé par l'apprenti lecteur pour comprendre le système alphabétique et son fonctionnement dans la mesure où ils servent d'agents médiateurs entre le patron orthographique et la représentation sémantique (Ecalle & Magnan, 2002). Pour apprendre à lire, l'enfant doit donc apprendre à relier les unités orales à leurs transcriptions graphiques, ce qui nécessite de développer des connaissances sur les liens entre les segments du langage parlé et les segments de l'écrit (Goswami & Bryant, 1990). Toutefois, la nature des unités utilisées dans ce traitement peuvent varier en fonction du niveau d'expertise des sujets.

Les travaux sur le rôle de la syllabe en tant qu'unité de traitement au cours de la reconnaissance des mots écrits se sont développés aussi bien en anglais qu'en français. Parmi les principales propositions émises, on peut citer les unités infra-syllabiques (Treiman, 1989) et les unités syllabiques (Colé, Magnan & Grainger, 1999). La syllabe est en effet décomposable en deux parties principales : l'attaque et la rime. L'attaque est constituée du groupe consonantique initial, par exemple /t/ dans « tard » et peut correspondre à une

consonne ou à un groupe de consonnes. La rime est constituée par le noyau vocalique et le coda consonantique terminal, par exemple la rime de « tard » est /ar/.

Mot pluri-syllabique	CALVICULE (/klavikul/							
Syllabes	**/kla/-/vi/-kul/**							
Attaque-rime	**/kl/-/a/**		**/v/-/i/**		**/k/-/ul/**			
Phonèmes	**/k/**	**/l/**	**/a/**	**/v/**	**/i/**	**/k/**	**/u/**	**/l/**

Figure 4 : Les différents niveaux de segmentation phonologique (Gombert & al., 2000)

Les résultats d'expériences menées par Treiman (1983 ; 1989) corroborent l'hypothèse du rôle central de la structure attaque-rime à l'oral et le fait que ces unités sont plus accessibles aux jeunes enfants anglais que les phonèmes. Ces enfants s'avèrent capables de décomposer en attaque-rime un mot monosyllabique (bloc) sans être capable de décomposer en deux phonèmes l'attaque /bl/ et la rime /oc/. La conscience d'unités intra-syllabiques telles que l'attaque et la rime pourrait bien apparaitre, au cours du développement de l'enfant, après la conscience des syllabes mais avant la conscience des phonèmes. L'enfant va découper la chaine graphique non pas en unités graphémiques correspondant aux phonèmes isolés mais en unités orthographiques correspondant aux unités intra-syllabiques (Treiman, 1992). Par exemple, pour lire le mot « *blast* », l'enfant pourrait utiliser des associations entre « *bl* » et /*bl*/ d'une part et entre « *ast* » et /*aest*/ d'autre part. Selon cette hypothèse, l'enfant peut mémoriser l'association entre « *bl* » et /*bl*/ et l'utilisera chaque fois qu'il rencontre « *bl* » au début d'un mot.

Le type d'unités utilisé dans le processus de segmentation semble lié aux caractéristiques phonologiques de la langue. Contrairement à l'anglais, le système phonologique français présente certaines caractéristiques favorisant l'utilisation d'unités syllabiques pour traiter le langage parlé (Kolinsky, Morais & Cluytens, 1995 ; Sprenger-Charolles & Casalis, 1996)[23]. La syllabe serait l'unité de segmentation fondamentale en français. Les travaux sur le rôle de l'unité syllabique dans les processus d'apprentissage de la lecture en français montrent que les jeunes lecteurs pourraient utiliser leur connaissance de la structure syllabique pour traiter le langage parlé (Ecalle & Magnan, 2007b).

[23] En anglais, la majorité des syllabes (60%) présente une structure fermée alors qu'en français, on trouve un plus grand nombre (76%) de syllabes ouvertes (55% de CV et 14% de CCV) (Sprenger-Charolles & Colé, 2003).

Colé, Magnan & Grainger (1999) ont fait l'hypothèse que les apprentis lecteurs francophones ont rapidement recours à des unités de type syllabique pour reconnaitre les mots écrits. Les sujets devaient décider le plus rapidement possible si une cible visuelle composée de deux ou de trois lettres (exemple PA ou PAR) était présente à l'initiale d'un mot présenté immédiatement après. Les cibles sont de structure phonologique CV ou CVC et correspondent ou non à la première syllabe des mots sur lesquels la tâche de détection est demandée. La mise en œuvre d'une procédure phonologique se traduit par un effet de comptabilité syllabique entre la cible et la syllabe initiale du mot. Les temps de détection sont plus rapides et les moyennes d'erreurs plus faibles lorsque la cible correspond à la première syllabe du mot (exemple : détecter PA dans PAROLE ou PAR dans PARDON) que lorsqu'elle ne correspond pas. L'effet de la longueur de la cible, c'est-à-dire des temps de détection qui ne sont fonction que de la longueur de la cible (détection plus rapide pour les cibles CV de deux lettres que pour les CVC de trois lettres), témoigne soit d'un traitement orthographique, lettre par lettre, soit d'un traitement phonologique effectué sur la base de correspondances grapho-phonémiques.

Les résultats de Sprenger-Charolles & Casalis (1996) soutiennent les hypothèses générales concernant les effets de la structure syllabique sur les performances en lecture. Le traitement syllabique mis en évidence dans cette étude varie en fonction du degré de complexité phonologique de la syllabe traitée. En effet, les enfants lisent et écrivent mieux les items à structure syllabique simple que ceux qui présentent une structure complexe, qu'il s'agisse des CCV/CVC ou des CVC/CVC. L'analyse des erreurs a également permis de voir que pratiquement tous les cas de suppression, en lecture comme en écriture, concernent la simplification d'une attaque complexe ou la disparition d'une coda. Les sujets ont donc tendance à réduire la syllabe à ses éléments principaux. Ces données révèlent que les erreurs sont bien phonologiques, elles traduisent la tendance des apprentis lecteurs francophones à simplifier les items polysyllabiques en les réduisant à une structure phonologique de type CV. Ecalle & Magnan (2007b) trouvent que la syllabe constitue une unité linguistique importante dans les processus phonologiques. Les enfants ont de meilleurs performances dans le traitement des unités en position initiale (syllabe, attaque). En revanche, la rime ne semble pas constituer une unité saillante en français.

L'ensemble des résultats présentés vont dans le sens de l'hypothèse selon laquelle la syllabe pourrait être une unité pertinenete de traitement au cours de l'apprentissage de la lecture. Ainsi, en début d'apprentissage, l'apprenti lecteur tente d'extraire de la chaine graphique du

mot des configurations orthographiques correspondant aux syllabes orales parce qu'il peut les percevoir et les produire. Le phonème, unité linguistique formelle de description de la langue, ne peut pas constituer, au tout début de l'apprentissage, une unité fonctionnelle de traitement (Ecalle & Magnan, 2002). Différents travaux ont mis en évidence les difficultés rencontrées par les enfants d'âge préscolaire pour formuler des jugements explicites concernant les phonèmes ou pour estimer le nombre de phonèmes dans un mot alors qu'ils réussissent des tâches portant sur des unités phonologiques plus grandes que le phonème (syllabe, rime). En effet, les phonèmes ne sont pas prononcés les uns à la suite des autres, mais en un seul geste articulatoire à l'intérieur d'une syllabe alors que la perception de la syllabe correspond à un acte articulatoire unitaire et facilement isolable (Alegria & Morais, 1989)[24]. Le développement de la conscience phonémique exige un degré d'abstraction important parce qu'il focalise l'attention de l'enfant sur des unités du langage parlé difficilement accessibles pour des raisons de coarticulation.

Les recherches menées en langue arabe sur les compétences phonologiques ont tenté de déterminer le rôle des unités telles que la syllabe et la rime dans la reconnaissance des mots écrits. Ammar (1997) a trouvé que les élèves bons lecteurs de 1ère, 2ème et 3ème années réussissent moins bien les tâches de segmentation phonémique que leurs homologues anglais ou français. Il explique les difficultés rencontrées par les lecteurs arabes par la situation de diglossie puisque les mots changent principalement de représentations phonologiques lors du passage de l'arabe standard à l'arabe dialectal.

Une étude de Ez-Zaher (2004, cité par Besse, 2007) confirme l'importance de la conscience phonémique en arabe au début de l'apprentissage de la lecture. La suppression de phonème apparaît être le meilleur prédicteur des habiletés de lecture une année plus tard, mais les compétences en segmentation syllabique apportent également une contribution non négligeable à l'explication de la variance en lecture. L'auteur explique ce résultat par le fait que les unités graphiques en arabe correspondent le plus souvent à la syllabe dû au marquage diacritique des voyelles, ce qui contraint le développement des capacités métaphonémiques et favorise, à l'écrit, une analyse en syllabes.

Concernant les unités intra-syllabique, les résultats des recherches réalisées sur la langue arabe s'accordent pour avancer que les enfants arabophones ne seraient pas sensibles à la rime

[24] Les systèmes d'écriture qui prennent la syllabe comme unité de base, le kana japonais en particulier, sont d'un accès plus facile pour l'enfant que les systèmes alphabétiques (Alegria & Morais, 1996).

puisque la structure bidimensionnelle du mot, formé d'une racine consonantique et d'un schème vocalique, ne serait pas basée sur la segmentation attaque-rime mais sur une décomposition morphologique en racine-shème (Ammar, 1997 ; 2003 ; Boukadida, 2008 ; Besse, 2008 ; Boukadida, Gombert & Maaouia, 2009). Dès le début de l'apprentissage de la lecture, les enfants auraient tendance à délaisser les voyelles au profit de la structure morphologique des mots.

1.2. Développement des capacités métaphonologiques

La conscience phonologique peut être définie comme « *la capacité qu'ont les sujets à prendre des distances par rapport au langage comme moyen de communication, à se représenter la parole comme une séquence d'unités discrètes, de dimension variée et à effectuer sur ces unités certaines opérations* » (Casalis & Lecoq, 1992 ; p. 198). C'est l'aptitude spécifique qui consiste à segmenter les mots en syllabes puis en phonèmes. Gombert (1992a) décrit cette capacité métaphonologique comme permettant d'identifier et de manipuler de façon intentionnelle les composantes phonologiques des unités linguistiques. Ces unités peuvent être syllabiques, intra-syllabiques ou phonémiques. Morais (1991) décrit quatre niveaux de conscience phonologique qui impliquent des demandes cognitives différentes, la conscience des chaînes phonologiques qui permet à l'enfant de repérer l'intonation, la capacité à repérer les rimes sans tenir compte de la signification, la conscience syllabique et enfin la conscience phonémique correspondant au plus haut niveau d'abstraction. La conscience phonémique semble constituer le niveau de compétence phonologique le plus lié au développement de la compétence lexique. L'apprentissage du code alphabétique repose, en effet, sur la capacité à décomposer les mots en phonèmes succédant à la capacité à diviser les mots en éléments syllabiques et intra-syllabiques.

Antony & al. (2003, cité par Ecalle & Magnan, 2007a) montrent que l'ordre d'acquisition des habiletés phonologiques dépend de deux facteurs, la taille de l'unité traitée et le type de tâche. Dans une étude menée auprès de 947 enfants de 2 à 6 ans, les auteurs ont examiné le niveau de sensibilité phonologique à l'aide de tâches d'association et de suppression, où il est demandé à l'enfant par exemple de former un mot avec 2 mots courts (*cow….boy - cowboy*), de supprimer une syllabe d'un mot pour former un autre ou de supprimer un phonème dans un mot pour trouver un mot ou un pseudo-mot. Le premier relève de la complexité linguistique : l'enfant maîtrise mieux d'abord le mot puis la syllabe, les unités infra syllabiques et enfin le

phonème. Le second facteur concerne la complexité de la tâche où on relève aussi 4 niveaux de difficulté : la détection de l'association, la détection de la suppression, la détection elle-même et la suppression elle-même.

Selon Muter, Hulme, Snowling & Taylor (1998), il existerait deux types d'habiletés phonologiques plus ou moins indépendantes, la segmentation phonémique et le traitement des rimes. Il semble que la conscience phonémique est importante pour la compréhension du principe alphabétique dans les premières étapes de la mise en place et de l'utilisation du traitement phonologique. Lorsque ce traitement s'automatise, l'enfant serait en mesure d'utiliser des unités de recodage plus larges que les phonèmes. Plusieurs niveaux d'habiletés phonologiques peuvent ainsi être envisagés. L'étude de ces habiletés nécessite, en conséquence, une distinction entre des processus au statut cognitif différent selon qu'ils sont accessibles ou non à la conscience (Ecalle & Magnan, 2003).

Pour rendre compte du développement des habiletés phonologiques, Gombert (1990) considère que le fonctionnement linguistique doit être étudié en distinguant les tâches qui exigent de la part de l'enfant une action délibérée et réfléchie, c'est-à-dire une attention réflexive sur le langage, de celles qui ne nécessitent qu'un traitement conscient. Ces deux comportements qui sont des versants des connaissances implicites et explicites mettent en évidence deux niveaux différents dans le contrôle cognitif qu'un individu peut exercer sur ses connaissances phonologiques : un niveau de traitement épiphonologique reposant sur une conscience implicite et un niveau de traitement métaphonologique reposant pour sa part sur une conscience explicite. Ces deux niveaux s'inscrivent dans un modèle de développement plus large (voir figure 4). Dans ce modèle, l'émergence des capacités métalinguistiques dépend à la fois de l'existence préalable de ces connaissances à un niveau épilinguistique et de l'intervention d'une demande externe qui oblige l'enfant à faire l'effort cognitif nécessaire à l'installation des possibiltés de contrôle conscient (Gombert, Bonjour & Marec-Breton, 2004).

Gombert (1990) remarque que les comportements épiphonologiques apparaissent à un âge très précoce chez l'enfant. Ce sont les jeux vocaux présents dès l'âge de trois ans ou les premières conduites de segmentation lorsque l'enfant doit supprimer la syllabe finale d'un mot par exemple. L'apparition des capacités métaphonologiques serait stimulée par l'enseignement formel de la langue écrite ou par un entraînement spécifique. Deux types de connaissances sont ainsi à l'œuvre dans la lecture, des connaissances implicites qui interviennent à l'insu du

lecteur et des connaissances explicites dont il contrôle consciemment l'intervention (Gombert & al., 2004).

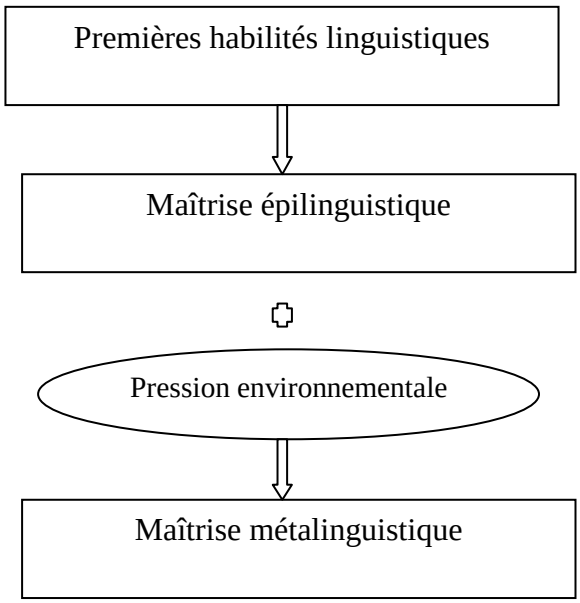


Figure 5 : Le développement métalinguistique (Gombert & al., 2004)

Bryant, Mac Lean, Bradley & Crossland (1990) ont tenté de vérifier l'hypothèse selon laquelle les habiletés épiphonologiques sont essentielles pour un futur apprentissage de la lecture en suivant longitudinalement 64 enfants. L'habileté épiphonologique a été mesurée à l'âge de 4 ans à l'aide de deux épreuves (détection de rimes et allitérations). A 6 ans, la compétence métaphonologique des enfants est évaluée par des épreuves de comptage et de suppression de phonèmes. Enfin, le niveau de maîtrise de la lecture est évalué à 7 ans à l'aide de deux tâches de lecture de mots. Les résultats montrent que le niveau de réussite aux épreuves épiphonologiques à 4 ans est relié à la performance en lecture deux ans plus tard ($r>.60$). Les variables détection de rimes et allitérations contribuent significativement à l'explication de la variance dans les épreuves de lecture. Les connaissances épiphonologiques semblent donc importantes car elles sont reliées à la future performance en lecture et se développent avant un réel apprentissage de la lecture (Marivain, 1992).

Ecalle, Magnan & Bouchafa (2002) ont examiné le développement des performances épiphonologiques et métaphonologiques auprès d'une population importante de grande section (GS) et de CP à différents moment de l'année. Un premier résultat classique confirme la plus forte progression des habiletés phonologiques sous l'effet de l'apprentissage de la

lecture-écriture au CP. Un second résultat confirme la facilité de traitement de la syllabe par rapport au phonème. Pour les syllabes aussi bien que pour les phonèmes, le niveau épiphonologique donne lieu à des performances significativement supérieures à celles du niveau métaphonologique seulement en début de GS. Cette tendance s'inverse en début de CP où les différences deviennent significatives au profit de l'extraction et l'identification d'unités phonologiques.

Dans une autre étude, Ecalle & Magnan (2003) ont pu montrer qu'avant l'instruction formelle de la lecture, les enfants font preuve de compétences phonologiques décalées selon le niveau de traitement impliqué. L'étude a porté sur un échantillon de 40 élèves de grande section de maternelle, non lecteurs, d'âge moyen 5, 7 ans. Deux tâches leur ont été proposées : une tâche épiphonologique où il s'agit de désigner deux mots (parmi quatre) partageant une unité phonologique et une tâche phonologique qui consiste à détecter l'unité commune à deux mots. Pour les deux tâches, trois types d'unités sont partagés par les mots cibles : une syllabe, une unité infra-syllabique ou un phonème, leur position variant dans le mot. Les mots étaient présentés à l'aide de supports imagés. Les résultats indiquent que les compétences dans une tâche épiphonologique sont supérieures à celles d'une tâche métaphonologique. D'autre part, l'effet du type d'unité traitée est clairement établi : les unités larges (syllabes, infra-syllabes) sont mieux traitées que les unités réduites. Dès lors, les auteurs proposent d'envisager un entraînement basé sur le traitement épiphonologique précédant et facilitant un entraînement impliquant des tâches métaphonologiques, ce qui pourrait contribuer à réduire les difficultés rencontrées par certains enfants au cours de l'apprentissage de la langue écrite.

Ecalle & Magnan (2007b) ont étudié le développement des habiletés phonologiques avant et pendant l'apprentissage de la lecture chez 34 apprentis lecteurs francophones. Les enfants ont été testé deux fois, en GSM et au CP à l'aide d'une épreuve de type épiphonologique (jugement de similarité) et de deux épreuves de type métaphonologique (extraction et substitution d'unités phonologiques). Les résultats indiquent que sous l'effet de l'instruction formelle de la langue écrite, les scores obtenus sur les unités phonémiques sont meilleurs que ceux obtenus sur les syllabes en particulier dans la tâche d'extraction d'unités phonologiques. Les données de cette recherche indiquent également que les enfants français possèdent une conscience excellente des syllabes avant l'apprentissage formel de la lecture. Les corrélations entre les performances en identification de mots écrits et les habiletés phonologiques confirment le lien puissant entre conscience phonémique et apprentissage de la lecture.

Les capacités métaphonologiques s'apparentent à une collection de capacités organisées dont certaines se manifestent bien avant l'apprentissage de la lecture, et d'autres sont plus tardives. Le passage de l'une à l'autre se fait dans un premier temps sur la base d'une prise de conscience de ses propres connaissances implicites puis sur la possibilité de les utiliser délibérément (Gombert & al., 2004). Cette prise de conscience n'émergerait que sous l'influence de facteurs externes comme, par exemple, l'abord de l'écrit dans un système alphabétique (Gombert, 1992). Il apparaît, ainsi, que certaines connaissances dites « épiphonologiques » peuvent se manifester très tôt chez l'enfant et peuvent servir de prédicteurs pour l'apprentissage de la lecture (Ecalle & Magnan, 2007a). La manipulation des syllabes et des rimes, facilement identifiables seraient sous-tendue par l'acquisition du langage oral et se développerait alors de façon presque naturelle à partir de jeux phonologiques (Gombert & Colé, 2000). En revanche, les connaissances métaphonémiques ne peuvent se manifester qu'après l'apprentissage de la lecture dans un système alphabétique. Se pose donc la question de la nature de la relation entre conscience phonologique et apprentissage de la lecture.

1.3. Apprentissage de la lecture et conscience phonologique : quels liens ?

De nombreux travaux établissent aujourd'hui, avec certitude, l'importance de la maîtrise métaphonologique au début de l'apprentissage de la lecture (Ecalle & Magnan, 2007a ; 2007b ; Ecalle, Magnan & Bouchafa, 2002 ; Gombert, 2004, 2003, 1990 ; Sprenger-Charolles & Casalis, 1996). Des liens très forts entre les performances des enfants dans des tâches portant sur les phonèmes et leur progrès dans l'apprentissage de la lecture ont été mis en évidence (Demont, 1994, 2003 ; Demont & Ecalle, 2007 ; Ecalle & Magnan, 2003 ; Sprenger-Charolles & Casalis, 1996 ; Alegria & Morais, 1989, Mann, 1989).

Les études longitudinales et prédictives ont mis en évidence l'existence d'une corrélation entre des tâches métaphonologiques réalisées avant les débuts de l'enseignement formel de la lecture et les résultats obtenus par les enfants dans son apprentissage. En effet, ces études ont montré que le niveau de conscience phonologique de l'enfant en maternelle constitue un des meilleurs prédicteurs de son niveau de lecture ultérieur (Shatsneider & al., 2004 ; Muetr & al., 2004 ; Demont & Gombert, 1996 ; Lecoq, 1991). Plus précisément, ces études montrent que seule la conscience phonémique par rapport à la conscience des unités plus larges telles la rime et la syllabe explique le mieux les performances en lecture. En plus, les études

longitudinales portant sur des enfants dyslexiques ont permis de relever un déficit d'analyse phonémiques chez de futurs dyslexiques avant même l'apprentissage de la lecture (Wimmer, 2003 ; Sprenger-Charolles, Colé, Lacert & Serniclaes, 2000). L'étude de Byrne, Freebody & Gates (1992) montre que les compétences phonologiques précoces, telles que révélées par la lecture de pseudo-mots, prédisent les compétences ultérieures en identification de mots écrits. Les auteurs ont suivi pendant un an des élèves scolarisés en 2ème et 3ème année du primaire. A l'intérieur de chaque niveau scolaire, les enfants ont été répartis en quatre groupes en fonction de leurs résultats à une épreuve de lecture de mots irréguliers et de pseudo-mots. Ceux qui manifestent de bonnes capacités en lecture de pseudo-mots mais pas en lecture de mots irréguliers sont appelés « phéniciens » et ceux qui présentent le profil inverse sont dénommés « chinois ». Les deux autres groupes comprennent les enfants qui ont réussi ou échoué dans les deux épreuves (bons lecteurs et faibles lecteurs). Les résultats observés un an plus tard en lecture de mots irréguliers montrent que « les phéniciens » et les « bons lecteurs » sont les seules qui améliorent leurs performances entre les sessions. Ces données suggèrent que les compétences phonologiques possèdent un fort pouvoir génératif puisque ce sont les seules qui prédisent les compétences ultérieures en identification des mots écrits. Le développement de la conscience phonologique serait un pré-requis de l'apprentissage de la lecture et favoriserait cet apprentissage. La conscience phonémique est, en effet, l'un des facteurs les plus importants de la réussite en lecture. La capacité de segmenter les mots en phonèmes est nécessaire pour que les enfants soient en mesure de découvrir les correspondances entre les éléments du langage parlé et écrit (CGP). La connaissance de ces CGP est à son tour nécessaire pour reconnaître les mots rencontrés pour la première fois et pour parvenir à l'acquisition de la vitesse et de l'automaticité de l'identification des mots (Gough & Hillinger, 1980 ; Jorm, Share, MacLean & Matthiews, 1984 ; Stanovich, 1986).

Par ailleurs, un certain nombre de données témoignent plutôt du fait que l'apprentissage de la lecture dans un système alphabétique constitue un déterminent majeur de l'émergence de la conscience phonologique. Cette position est essentiellement justifiée par les études ayant montré que les enfants prélecteurs aussi bien que les adultes illettrés ont une énorme difficulté pour segmenter les mots en phonème même si la segmentation syllabique et la détection de la rime sont relativement faciles (Alegria & Morais, 1979 ; Morais, Cary, Alegria & Bertelson, 1979). Des travaux similaires portant sur deux groupes de lecteurs chinois pratiquant une écriture alphabétique et non alphabétique confirment ces conclusions (Read, Zhang, Nie & Ding, 1986). Les premiers ont appris à lire dans un système d'écriture alphabétique alors que

les seconds n'ont appris que les caractères logographiques traditionnels. Les résultats ont montré que seuls ceux qui avaient appris à lire un alphabet étaient capables de manipuler volontairement des phonèmes. Ce n'est donc pas la connaissance d'une langue écrite qui, en tant que telle, apparaît favoriser la réussite dans la manipulation d'unités phonémiques mais bien la maîtrise du code alphabétique. Gombert (1992a) souligne de la même façon que la conscience phonémique ne peut réellement se développer que dans un contexte d'instruction du principe alphabétique, « *ce n'est pas l'apprentissage de la lecture en général, mais l'apprentissage de la lecture d'une langue alphabétique qui est liée à la capacité métaphonologique* » (p. 114). Certains auteurs ont tenté de séparer les effets de l'âge et de l'apprentissage dans le développement des capacités métaphonologiques (Bentin, Hammer & Cahan, 1991). Ils constatent que l'effet net de la scolarité sur les performances métaphonologiques s'élève environ à 32% alors que celui de l'âge se réduit à 9%. L'effet de scolarité est donc quatre fois plus important que celui de l'âge.

Il existe actuellement un consensus pour admettre que la conscience phonologique et l'apprentissage de la lecture se développent en interaction et se renforcent mutuellement (Demont & Gombert, 2007 ; Perfetti, 1989 ; Perfetti, Beck, Bell & Hughes, 1987). D'une part, la conscience phonologique faciliterait l'apprentissage de la lecture et d'autre part, les progrès en lecture semblent favoriser le développement des capacités métaphonologiques. Dans le premier cas, la conscience phonologique apparaît comme étant la cause de la réussite ou de l'échec dans l'acquisition de la lecture. Dans le second, elle apparaît comme la conséquence de cette acquisition. De fait, la conscience phonologique apparaît comme bidirectionnelle, à la fois cause et conséquence de l'apprentissage de la langue écrite et constitue un bon prédicteur de la réussite en lecture (Mann , 1989). Selon Demont & Gombert (2007), le lien entre développement des compétences métaphonologiques et apprentissage de la lecture s'apparenterait davantage à une relation circulaire.

En ce qui concerne les langues sémitiques, l'étude de Bentin & Leshem (1993) semble montrer que l'enseignement de la lecture en hébreu favorise l'apparition de la conscience phonémique chez les enfants dès l'âge de 6 et 7 ans. En outre, ce même apprentissage est amélioré par un entraînement à la segmentation phonémique. Selon les auteurs, il y aurait en hébreu une relation de réciprocité entre les deux compétences. En arabe, cette relation semble problématique dans la mesure où la langue possède une écriture vocalisée basée sur des correspondances grapho-phonologiques strictes, mais dont la structure serait plutôt morphologique que phonologique.

Les recherches menées en langue arabe sur les compétences phonologiques confirment le rôle primordial de la conscience phonémique lors de la reconnaissance des mots écrits. Les travaux de Ammar (1997) ont montré, en effet, que les performances au niveau des traitements phonémiques (tâches de comptage, de segmentation et de suppression de phonèmes) sont fortement corrélées avec les progrès en lecture. De même, les comparaisons bons lecteurs/faibles lecteurs (Ben Dror, Bentin & Frost, 1996 ; Ammar, 1997 ; Belajouza, 2003) ont mis en évidence des déficits importants au niveau des compétences phonologiques, essentiellement métaphonémiques, chez les faibles lecteurs. L'apprenti lecteur arabophone s'appuie, comme l'apprenti lecteur en langues à écriture latine, sur ses connaissances du fonctionnement de la langue orale pour mettre en place la stratégie alphabétique.

Abu Rabia (1995) a conduit une étude sur 143 apprentis lecteurs arabophones âgés de 8 à 11 ans. Les résultats montrent que, comme en anglais, il existe une relation étroite entre l'identification des mots en arabe et les compétences phonologiques, les connaissances syntaxiques et sémantiques et la mémoire à court terme. Les compétences phonologiques et plus particulièrement la connaissance des règles de correspondances grapho-phonologiques jouent un rôle essentiel dans l'apprentissage de la lecture en arabe et peuvent faciliter le décodage rapide des mots. Ceci est du aux correspondances graphophonologiques (CGP) strictes en arabe. Les faibles lecteurs montrent des performances faibles au niveau des compétences phonologiques et sémantiques.

En langue arabe, la nature de la relation entre la conscience phonologique et l'apprentissage de la lecturen n'a pas été formellement précisée, notamment si l'on tient compte de la structure morphologique de la langue et surtout de l'existence de deux codes : le code oral (celui de l'arabe dialectal, langue orale) et le code écrit (celui de l'arabe standard, langue écrite et enseignée). Les résultats de Ammar (1997) ont montré que les enfants qui n'ont pas abordé l'apprentissage formel de la lecture se sont montrés performants dans certaines tâches d'analyse syllabique, ce qui relève d'un comportement épiphonologique et leurs performances s'améliorent au fur et à mesure qu'ils avancent dans la scolarité, ce qui confirme l'idée selon laquelle l'apprentissage de la lecture favorise l'émergence de la conscience phonologique. De même, les enfants parviennent à procéder à des analyses phonologiques quelle que soit la nature du mot à manipuler : en arabe standard, en arabe dialectal ou pseudo-mots, ce qui confirme l'existence d'une relation de causalité réciproque entre l'apprentissage de la lecture et l'acquisition de la conscience phonologique. Conformément à la langue anglaise, la conscience phonologique joue un rôle fondamental dans l'apprentissage de la lecture en arabe.

De plus, les habiletés phonologiques semblent avoir également le statut de cause et de conséquence de l'acquisition de la lecture en langue arabe.

En résumé, ces différents travaux font apparaître l'importance des habiletés métalinguistiques, notamment métaphonologiques, dans l'apprentissage de la lecture. Toutefois, en plus de la nécessité de la maîtrise du principe alphabétique, cet apprentissage mobilise également des connaissances morphologiques. La capacité à réfléchir à la langue nécessaire à l'apprentissage de la lecture doit aussi s'exercer sur les éléments qui constituent la structure morphologique du mot.

2. Les connaissances morphologiques

L'importance de la maîtrise des CGP pour l'apprentissage de la lecture est clairement démontrée, l'ensemble des recherches en rend compte. Cependant, il est important de préciser que les compétences phonologiques ne sont pas les seuls pouvant exercer une influence sur l'apprentissage de la lecture. D'autres recherches évoquent, en effet, l'importance des connaissances morphologiques dans l'acquisition des procédures d'identification des mots écrits (Boukadida, 2008 ; Carlisle, 1995 ; Gombert, 2003 ; Marec-Breton, 2009 ; 2003), mais il semblerait que leur rôle soit moins prépondérant que celui des capacités phonologiques (Marec-Breton, 2003).

La morphologie est l'étude de la structure des mots et de l'organisation des unités de sens qui composent les mots (Gombert & al., 2000). L'élève peut décomposer un mot en unités morphologiques, c'est à dire en unités de signification appelées morphèmes, définis comme les plus petits segments qui véhiculent du sens. Ainsi, on peut distinguer les mots morphologiquement simples (composés d'un seul morphème) des mots morphologiquement complexes (composés d'au moins deux morphèmes. Par exemple, le mot "chat" constitue à lui seul un morphème. En revanche, le mot "chaton" est composé de deux morphèmes : la base chat- et le suffixe -on ; le premier désignant un animal particulier et le second un diminutif qui permet d'aboutir à la signification "petit chat". La combinaison de morphèmes constitue donc un mécanisme de création de mots extrêmement puissant.

Le champ de la morphologie s'organise en deux grandes sphères. D'une part, la morphologie flexionnelle dont le domaine est constituée par les variations formelles que subissent les mots en rapport avec leur fonction syntaxique. D'autre part, la morphologie dérivationnelle qui

s'intéresse à la construction des mots et possède une fonction sémantique importante. La maîtrise des règles dérivationnelles s'acquiert plus lentement et plus tardivement que celle des règles flexionnelles et il faut parfois attendre le lycée pour que le sens de certains suffixes dérivationnels soit maîtrisé (Colé & Fayol, 2000). Les études conduites sur le développement des connaissances morphologiques s'intéressent essentiellement à la capacité de l'enfant à réfléchir et à manipuler explicitement la structure morphémique des mots. Elles reposent le plus souvent sur des tâches de production et de jugement des relations morphologiques entre les mots. La question qui se pose est la suivante : les connaissances morphologiques de l'enfant peuvent-elles jouer un rôle précoce dans l'apprentissage de la lecture et plus particulièrement dans l'élaboration des procédures de reconnaissance de mots écrits ?

2.1. La place de la morphologie dans l'apprentissage de la lecture

Comparativement aux travaux sur le développelement de la conscience phonologique, peu de recherches se sont intéressées à la conscience morphologique. La plupart des recherches sur l'apprentissage de la lecture considèrent que le traitement morphologique des mots ne joue un rôle dans l'apprentissage de la lecture qu'après plusieurs années de lecture puisque les connaissances morphologiques concernent des éléments véhiculant des informations syntaxiques et sémantiques. L'analyse morphologique n'apparaît pas primordiale par rapport à l'analyse phonémique et aurait un rôle plus tardif (Leybeart & Alegria, 1995). Ainsi, la majorité des modèles d'apprentissage de la lecture réduisent l'installation de l'habileté de reconnaissance des mots écrits au décodage et à l'assemblage phonologique et ignorent presque totalement la dimension morphologique des mots. Par exemple, pour Frith (1985), la structure morphologique des mots ne serait utilisée dans la lecture qu'à partir du moment où l'élève maîtrise le décodage graphophonologique c'est-à-dire au stade orthographique. Ce recours à la morphologie serait essentiellement destiné à pallier l'irrégularité de certains mots qui ne peuvent pas être lus par simple conversion des graphèmes en phonèmes. De fait, une grande majorité des recherches concernant l'impact de la morphologie sur l'apprentissage de la lecture ne s'intéresse à cette question qu'à partir de la troisième année de l'enseignement primaire. L'hypothèse de la plupart des auteurs est que la morphologie ne peut intervenir dans les traitements opérés par les lecteurs qu'à partir du moment où les enfants maîtrisent le code alphabétique.

En français, Lecoq & al. (1996) soulignent que la notion de famille morphologique est loin d'être pleinement acquise au début de l'école primaire. Ils proposent à des enfants de choisir parmi trois mots : un dérivé du mot cible (*pommier* dérivé de *pomme*), un mot relié sémantiquement (poire) et un mot relié orthographiquement (*pommade*), celui qui constitue le dérivé du mot cible (*pomme*). Les auteurs observent 48% de réussite à cette tâche en grande section maternelle (GSM) et constatent par ailleurs que les performances augmentent entre la GSM et le cours préparatoire (CP) mais qu'elles ne diffèrent pas entre le CP et le CE1. Les travaux menés en italien par Burani, Marcolini & Stella (2002) montrent également que dès la troisième année, 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 construits par l'association illégale d'une base et d'un suffixe dérivationnel (*Donnista*) par rapport à ceux obtenus par des pseudo-mots ne contenant pas d'affixe. Les résultats révèlent que les pseudo-mots affixés sont lus plus vite et avec moins d'erreurs que les pseudo-mots sans structure morphologique.

Dans l'ensemble, ces travaux montrent que les enfants tirent partie des éléments morphémiques composant les mots au cours de la lecture et ce, quelle que soit la langue testée. Toutefois, aucune de ces études ne s'intéresse aux connaissances morphologiques au tout début de l'apprentissage de la lecture. Des travaux récents font l'hypothèse que dès les premières étapes de l'apprentissage de la lecture, avant donc la maîtrise complète des CGP, les enfants disposent de connaissances morphologiques implicites et explicites non négligeables (Colé, Marec-Breton, Royer & Gombert, 2003 ; Marec-Breton, 2009 ; 2003 ; Marec-Breton, Gombert & Colé, 2005 ; Gombert, 2003). Ces études ont, en effet, prouvé que les enfants dès 3 - 4 ans sont sensibles aux unités de sens véhiculées par la langue.

Colé & al. (2003) ont mis en évidence l'existence d'une sensibilité à la morphologie des mots oraux avant l'apprentissage de la lecture. Il a été ainsi montré que des enfants pré-lecteurs de 5 ans trouvent que les pseudo-mots oraux ayant une structure préfixe + base (par exemple, "préfader" / préfade/) ressemblent plus à des vrais mots que les pseudo-mots construits avec le même matériel phonologique (par exemple, "pradéfer" / pradefe/) mais n'ayant pas cette structure morphologique. De plus, les auteurs constatent que les lecteurs débutants (1ère année et 2ème année) lisent mieux les mots préfixés (*retomber*) que les mots pseudo-préfixés (*renifler*) et les pseudo-mots construits sur le principe préfixe + vrai mot (*remourir*) que ceux construits sur le principe préfixe + pseudo-mot (*recounir*). Ces résultats suggèrent que, dès le début de l'apprentissage de la lecture, il existe une prise en compte de la structure

morphologique des mots. L'apprentissage de la lecture accroît cette sensibilité et provoque l'apparition de connaissances morphologiques conscientes.

Les résultats obtenus dans la série d'expériences rapportée par Marec-Breton (2003) suggèrent que, dès le CP, alors que les enfants ne maîtrisent pas encore totalement le code alphabétique, des traitements morphologiques sont impliqués dans la reconnaissance des mots écrits. En effet, les enfants lisent mieux (plus vite et avec moins d'erreurs) des items affixés (*dégarer*) que des items pseudo-affixés (*démaner*) et savent reconnaitre deux mots de la même famille. Ainsi, dans une tâche de jugement de relations morphologiques (par exemple : de manger et boire, lequel est de la même famille que manger), les enfants répondent au moins à neuf items sur les dix proposés. Ils sont capables également de segmenter des mots afin d'en restituer les morphèmes constituants (par exemple, identifier un mot caché au sein d'un autre) et enfin, ils peuvent utiliser la structure morphologique des mots qu'ils rencontrent pour la première fois afin d'en découvrir la signification. Ces résultats valident donc, en français, l'existence d'une analyse morphologique dès les premières étapes de l'apprentissage de la lecture. La prise en compte de la structure morphologique des mots aide le lecteur débutant dans le traitement des mots écrits. Selon Gombert & al. (2000), les performances d'analyse morphologique sont très faibles et variables selon les tâches. Néanmoins, elles augmentent notablement de la grande section au CP (30% de réussite en moyenne) mais varient peu entre le CP et le CE1 (50%).

Colé, Royer, Leuwers & Casalis (2004) ont examiné le développement des connaissances morphologiques implicites et explicites d'enfants d'une classe de CP qui ont été suivis longitudinalement pendant un an. Ces connaissances ont été mises en relation avec le niveau de lecture. Les résultats observés au CP montrent que seules les performances obtenues à la tâche de jugement de relation de mots permettent de distinguer les niveaux de lecture. Ainsi, les enfants ayant obtenus les meilleures performances à cette tâche sont égalment ceux qui obtiennent les meilleurs scores de lecture. Pour ce qui concerne les tâches morphologiques plus explicites, les auteurs constatent que les performances à la tâche d'extraction de base au CP ne permettent pas de distinguer ni les niveaux de lecture ni les niveaux de vocabulaire. Ces données suggèrent que les connaissances morphologiques explicites ne seraient pas utilisées au cours de la lecture au CP et il faut attendre le CE1 pour qu'elles soient bien réussies. Les auteurs proposent ainsi un modèle qui permet d'envisager, dès le début de l'acquisition des règles de correspondances grapho-phonologiques, une intervention progressive des connaissances morphologiques (figure 6).

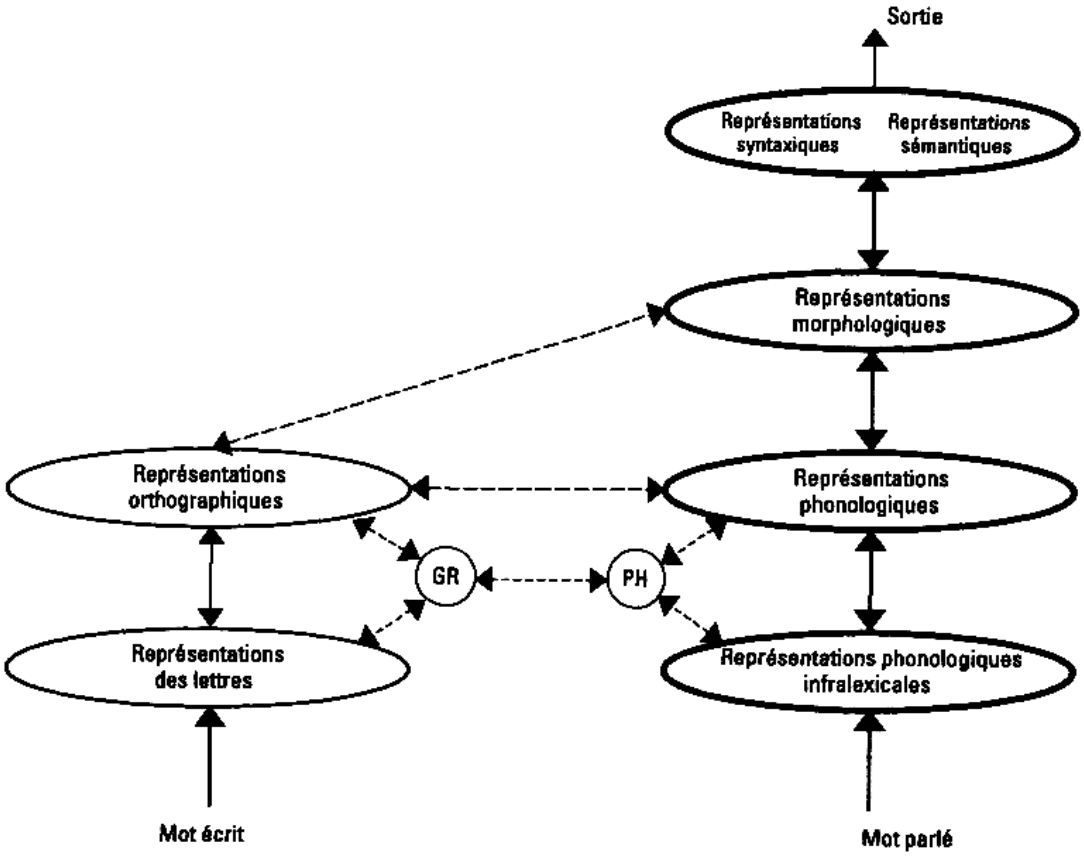


Figure 6 : Modèle théorique proposant une intervention précoce de la dimension morphologique dans l'apprentissage de la lecture (d'après Colé & al., 2004)

Ce modèle considère que la tâche principale de l'apprenti lecteur consiste, d'une part, à élaborer des représentations orthographiques des formes globales des mots et des représentations infra-lexicales (lettres) en liaison avec les représentations phonologiques. Sous l'action de l'apprentissage explicite des correspondances grapho-phonologiques et de l'exposition répétée aux mots écrits, l'apprenti lecteur développerait alors des connexions entre les représentations des lettres et les représentations d'unités phonologiques plus larges telles que les syllabes. Les représentations morphologiques sont activées dès que l'apprenti lecteur est capable d'utiliser la médiation phonologique. Le développement des connaissances morphologiques et l'acquisition de l'orthographe permettrait de créer un lien direct entre représentation orthographique et représentation morphologique.

En anglais, l'existence de connaissances morphologiques précoces est aussi démontrée par Laxon, Richard & Coltheart (1992). Dans cette étude, les auteurs soumettent des enfants âgés de 7 à 10 ans à une tâche de lecture à haute voix de mots et de pseudo-mots variant par leur

structure morphologique. Quatre types d'items sont proposés : des mots affixés (*dancer*), des mots pseudo-affixés (*dinner*), des pseudo-mots construits à partir d'une base existante et d'un affixe (*fooder* de *food*) et des pseudo-mots construits à partir d'une fausse racine et d'un affixe (*gonder*). Les résultats indiquent que la lecture à haute voix de mots suffixés occasionne significativement moins d'erreurs que celle des pseudo-mots et que les enfants lisent plus précisément les pseudo-mots composés d'une vraie racine et d'un suffixes que les pseudo-mots sans éléments morphémiques. Ces résultats sont obtenus à la fois avec les bons et les moins bons lecteurs. Les auteurs constatent également que la lecture des items finissant par *-er-* sont lus avec plus de précision que ceux finissant par *-ed-*. Cette différence serait due à la complexité des règles phonologiques des mots se terminant après *-ed-*. Les résultats de cette expérience confirment que l'information morphologique des mots est prise en compte au cours de la lecture et suggèrent que l'utilisation de cette information est sous la dépendance des règles phonologiques.

En langue arabe, Boukadida, Gombert & Maaouia (2009) ont tenté de vérifier que les apprentis lecteurs tunisiens utilisent les informations morphologiques dans la reconnaissance des mots écrits et ceci dès les premières étapes de l'apprentissage de la lecture. L'étude a porté sur des élèves tunisiens évalués à deux reprise, une premois fois à la fin de la 1ère année et une deuxième fois en à la fin de la 2ème année, grâce à des tâches phonologiques et morphologiques. Les résultats obtenus montrent que les élèves possèdent des connaissances morphologiques et ceci dès la 1ère année primaire. A ce niveau, ils commencent à intégrer la notion morphologique avec un taux de réussite de 60% pour les bons lecteurs et 45% pour les faibles lecteurs.

Avant l'apprentissage de la lecture, l'enfant dispose ainsi de connaissances morphologiques non négligeables, bien qu'elles soient moins développées que les connaissances phonologiques (Colé & Fayol, 2000 ; Marec-Breton, 2003). C'est sur la base de ces connaissances morphologqiues implicites que se développe la conscience morphologique (Gombert, 2003 ; Marec-Breton, 2003).

2.2. Rôle des connaissances morphologiques dans l'apprentissage de la lecture

Les premiers travaux sur le rôle des unités morphologiques dans l'apprentissage de la lecture ont été principalement réalisés sur la langue anglaise. Ces travaux postulent, en effet, que les

processus engagés lors de l'identification des mots sont sensibles à la structure morphologique interne des mots. Quelques auteurs s'intéressent depuis quelques années à une autre hypothèse qui fait état d'un déficit morphologique chez des sujets présentant un trouble d'acquisition de la lecture (Abu Rabia, 2007 ; Abu Rabia & al., 2003 ; Ben Dror & al., 1995 ; Elbro & Arnback, 1996 ; Lecoq & al., 1996 ; Shankweiler & al., 1995). Les travaux sur la dyslexie (Casalis, 1995 ; Elbro & Arnbak, 1996) ont montré que les enfants présentant un déficit phonologique important développent des stratégies compensatoires s'appuyant sur l'utilisation de l'information morphologique. La connaissance de l'organisation morphologique du mot faciliterait l'accès au lexique pour les lecteurs en difficulté. Par conséquent, Sabatier & al. (2003) émettent l'hypothèse que, tout comme il y aurait une conscience phonologique, il existerait également une conscience morphologique.

La conscience morphologique, ou compétence métamorphologique, est définies par Carlisle (1995) comme la conscience qu'a l'enfant de la structure morphémique des mots et sa capacité à réfléchir et à manipuler explicitement cette structure[25]. De nombreuses recherches évoquent, dès le début de l'apprentissage de la lecture, un lien entre le niveau de conscience morphologique et le niveau de lecture (Boukadida, 2008 en arabe ; Marec-breton, 2003 ; Casalis & Louis-alexandre, 2000 en français ; Carlisle & Nomanbhoy, 1993 en anglais). Ces recherches concernant les liens entre conscience morphologique et apprentissage de la lecture se sont focalisées, pour une grande part, sur le domaine de la morphologie dérivationnelle. Carlisle (1995) constate ainsi que les connaissances sur la morphologie dérivationnelle constituent un meilleur prédicteur de la réussite en lecture que celles concernant la morphologie flexionnelle puisqu'elles intègrent différentes dimensions linguistiques (phonologique, sémantique et syntaxique).

La recherche de Carlisle & Nomanbhoy (1993) montre que les connaissances métamorphologiques sont impliquées dans la lecture de mots dès la première année. Deux tâches métamorphologiques ont été utilisées, une tâche de jugement de liens morphologiques et une tâche de production de mots dérivés et fléchis. Les résultats indiquent que les performances des enfants de CP aux épreuves de conscience phonologique et morphologique contribuent toutes deux de façon significative à la variance de la lecture de mots isolés, même si la contribution apportée par la première est plus importante (respectivement 33 et 4%). Les

[25] Il convient de distinguer entre conscience morphologique et conscience syntaxique. La première est d'abord une conscience de la structure des mots. La seconde, au contraire, concerne directement la phrase mais ne concerne qu'indirectement le mot.

auteurs observent également que le niveau de conscience phonologique contribue de façon significative à la réussite aux épreuves de conscience morphologique.

Carlisle (1995) s'intéresse à l'aspect prédictif de la conscience phonologique et de la conscience morphologique évaluées en grande section et en CP sur le niveau en lecture-compréhension en CE1. Elle propose deux tâches aux enfants : une tâche de production où les sujets doivent compléter une phrase à partir d'une base qui leur est fournie (dance, a person who dance is a …) et une tâche de compréhension où ils doivent dire si la phrase qu'ils entendent a un sens (a person who teach is a teacher, a man who makes dolls is a dollar). Carlisle constate que les performances des enfants au niveau du CP aux épreuves de conscience morphologique contribuent à celles obtenus en lecture au niveau du CE1, ce qui confirme l'existence d'une relation entre conscience morphologique et apprentissage de la lecture. D'autre part, elle montre que la réussite à ces épreuves de conscience morphologique dépend du niveau phonologique des enfants. Ainsi, les enfants ayant de bonnes performances en analyse phonémique (testée par une tâche de suppression de phonèmes) réussissent mieux l'épreuve de production de mots dérivés. De la même façon, Tornéus (1990), en réalisant une étude longitudinale de la fin de la maternelle au niveau CE2, montre que les connaissances métamorphologiques mesurées en GSM et en CP sont un bon prédicteur des habiletés de lecture en CE1 et en CE2. Les études conduites avec des élèves plus âgés rapportent des résultats similaires. Ainsi, Fowler & Liberman (1995) observent avec des enfants de la 2ème à la 4ème année, des corrélations significatives entre les performances aux tâches morphologiques et celles de lecture de mots et de pseudo-mots.

En français, l'étude longitudinale de la GSM à la 2ème année conduite par Casalis & Louis-Alexandre (2000) montre que les performances précoces en morphologie dérivationnelle n'ont qu'un impact tardif sur la lecture. Les auteurs n'observent pas de contribution des tâches morphologiques implicites à la variance en lecture de mots isolés au CP alors que la tâche de suppression de phonèmes rend compte de 37,3% de cette variance. En revanche, au CE1, les résultats montrent que la tâche de production morphologique rend compte de 16% de la variance en lecture de mots isolés et la tâche de segmentation morphologique de mots complexes en base et affixe explique 24% de la variance en compréhension écrite.

Les résultats obtenus par Marec-Breton (2003) confirment l'existence d'un lien entre les connaissances métamorphologiques et la lecture. Les enfants obtenant les scores les plus élevés à la tâche de production de néologisme sont aussi ceux obtenant les scores les plus

élevés aux épreuves de lecture. Par exemple, dans une tâche évaluant le niveau de sensibilité morphologique, les enfants doivent produire un néologisme possible en appliquant une règle de construction morphologique, indiquée par une phrase contexte, sur le mot qui leur est donné. Ainsi, les enfants peuvent compléter la phrase contexte : *lampe, une petite lampe est une*… par le néologisme *lampette*. Les résultats montrent que les performances obtenues aux différentes tâches d'évaluation de la conscience morphologique sont corrélées significativement avec les scores en lecture. Par ailleurs, cette participation des connaissances morphologiques à la lecture s'observe dès le CP. A ce niveau, les performances obtenues à la tâche de production de néologisme expliquent 15% de la variance observée à la tâche de lecture en une minute.

Dans une étude plus récente, Marec-Breton, Gombert & Colé (2005) montrent la présence d'effets de la structure morphologique sur l'excactitude et le temps de lecture chez des enfants de niveau CP et CE1. Les résultats indiquent que la structure morphologique des items intervient lors de la lecture de mots non familiers. Elle constitue ainsi pour les enfants une aide notable à la lecture de mots nouveaux mais ne serait pas indispensable pour la lecture des mots connus. De la même façon, le recours à une analyse morphologique est plus systématique et plus précoce pour les items préfixés que les items suffixés. Ces résultats suggèrent que dans les premières étapes de l'apprentissage de la lecture, l'extraction d'information morphologique pourrait faciliter la lecture de certains items même si le décodage n'est pas complètement maîtrisé.

Le rôle des compétences morphologiques dans l'apprentissage de la lecture en langues sémitiques, telles que l'arabe et l'hébreu, a été jusqu'à maintenant peu exploré. Ces écritures accordent une place plus importante à la dimension sémantique et notamment à la morphologie. L'étude de Ben Dror, Bentin & Frost (1996) avait pour objectif de comparer les performances sémantiques, phonologiques et morphologiques d'un groupe de faibles lecteurs en hébreu avec deux groupes d'enfants normo-lecteurs dont l'un est apparié sur la base de l'âge et le second apparié sur la base de l'étendue du vocabulaire. Les résultats des faibles lecteurs ont été inférieurs à ceux du groupe contrôle lorsqu'il s'agissait d'assigner les mots à des catégories, d'identifier le premier phonème des mots prononcés et d'évaluer les relations morphologiques entre les couples de mots. Leurs difficultés ont été encore plus évidentes dans le domaine morphologique.

En arabe, le rôle des compétences morphologiques dans l'apprentissage de la lecture revêt un intérêt particulier vu la structure bidimensionnelle du mot. En langue arabe, le mot est constitué d'une racine composée seulement de consonnes. Les voyelles sont représentées par des signes diacritiques juxtaposées par rapport au squelette consonantique. Les résultats de Ammar (1997) ont mis en évidence la nécessité d'une analyse morphologique préalable au recodage grapho-phonologique. La structure morphologique de l'arabe, fondée sur la racine consonantique, laisse supposer que les mots sont plus facilement décomposables en racine consonantique et morphèmes vocaliques. Dans une expérience plus récente, Ammar (2003) confirme le traitement séquentiel des consonnes et des voyelles au cours de l'identification des mots écrits en langue arabe. L'étude a porté sur 72 élèves tunisiens de 2ème, 4ème et 6ème année de l'enseignement primaire. Dans une épreuve de lecture à haute voix de pseudo-mots vocalisés, Ammar observe que les erreurs de prononciation concernent exclusivement les voyelles, la prononciation des consonnes est le plus souvent correcte. Il semble ainsi que l'identification de mots en arabe s'appuie fondamentalement sur le squelette consonantique correspondant à la racine du mot et nécessite donc une analyse morphologique préalable. Dans le même sens, la recherche de Belajouza (2003) sur des enfants tunisiens confirme l'hypothèse du rôle des connaissances morphologiques. Elle propose quatre tâches à un groupe d'élèves normolecteurs en début de 2ème année ou en début de 3ème année et à un groupe d'élèves plus avancés dans la scolarité (5ème ou 6ème année) mais ayant de graves difficultés de décodage : une tâche d'identification de mots en contexte imagé, une tâche de lecture à voix haute de mots et de pseudo-mots, une épreuve de suppression et une épreuve de reconnaissance de liens morphologiques. Les résultats montrent que les faibles lecteurs présentent des déficits phonologiques et morphologiques importants.

Boukadida (2008) a tenté d'étudier le rôle des connaissances phonologiques et morphologiques dans l'apprentissage de la lecture en arabe. Les analyses de corrélation montrent l'existence de liens entre les connaissances métaphonologiques et métamorphologiques et la lecture. Les performances des élèves dans les tâches de morphologie dérivationnelle, avant le début de l'apprentissage de la lecture, sont même plus corrélées avec les performances en lecture une année plus tard que ne le sont les performances dans la manipulation des unités sonores. De plus, les analyses de régression ont montré que les compétences morphologiques en CP et en 1ère année, contribue de façon significative et indépendante à l'explication de la variance en lecture. L'ensemble des résultats confirme l'hypothèse d'un traitement morphologique lors de la lecture de mots en arabe et du rôle de la

racine dans la reconnaissance des mots. La décomposition des mots en racine et schème semble concerner les mots connus et les pseudo-mots.

Dans l'ensemble, ces données indiquent que le développement des habiletés de décodage nécessite également des compétences d'analyse morphologique pour les apprentis lecteurs en langue arabe. Il semble, par ailleurs, que le niveau de maîtrise des habiletés morphologiques, comme celui des habiletés métaphonologiques, soit corrélé avec le niveau des compétences en identification des mots et ceci dès le début de l'apprentissage.

3. La connaissance des lettres

La connaissance des lettres est une étape importante dans l'acquisition de la lecture dans un système d'écriture alphabétique (Colé & Valdois, 2007). Elle renvoie non seulement à l'identité visuelle des lettres mais fournit également un grand nombre d'indices permettant de déduire le son des lettres et d'établir ainsi les premières correspondances graphème-phonème (Treiman, 2006). La connaissance de la structure phonologique des noms des lettres facilite la compréhension du principe alphabétique dans la mesure où les enfants pré-lecteurs relient mots oraux et mots écrits en associant lettres et noms des lettres. Selon Ecalle & Magnan (2002), l'apprentissage des correspondances grapho-phonémiques est plus facile lorsque le nom de la lettre est acquis puisque l'établissement de la relation son-lettre peut s'appuyer sur le nom de la lettre contenant le phonème à mémoriser. Grâce au nom des lettres, les jeunes enfants peuvent progresser d'un traitement essentiellement visuo-sémantique des mots à un traitement linguistique préparant la voie à la procédure alphabétique (Biot-Chevrier, Ecalle & Magnan, 2008).

Dés les premières années de scolarisation, l'enfant est sensible aux fréquences des suites de lettres de la langue écrite. Ainsi, des connaissances orthographiques s'élaborent par apprentissage implicite sans que l'apprentissage des relations grapho-phonologiques soit impliqué. Au début, quelques unes de ces lettres suffisent pour identifier un mot (généralement une ou deux lettres initiales ou finales) puis s'impose la nécessité de toutes les lettres quel qu'en soit l'ordre. Enfin, apparaît la nécessité que les lettres composant le mot soient, non seulement toutes présentes, mais qu'elles soient dans un ordre strict (Bastien-Toniazzo, 2003). La connaissance de l'alphabet chez le jeune enfant montre l'existence d'un décalage entre la connaissance du nom des lettres et les connaissances relatives à leur valeur phonémique. Les premières étant acquises plus facilement et plus précocement chez les pré-

lecteurs (Ecalle, 2004). Le nom des lettres est indiqué très tôt aux enfants, généralement de manière implicite, dans de nombreuses situations éducatives (jeux télévisés, lecture d'abécédaires) alors que la valeur phonémique des lettres se développe en général sous l'effet de l'instruction formelle (apprentissage explicite).

Treiman, Weaterson & Berch (1994 ; cités par Ecalle, Magnan & Ramus, 2007) montrent que lorsque le nom de la lettre ne comporte pas sa valeur phonémique, cette dernière reste plus difficile d'accès. Les enfants d'âge préscolaire, en connaissant le nom de la lettre Y /*wai*/, lui affectent la valeur phonémique portée par la lettre W. La position du phonème dans le nom de la lettre est un autre facteur affectant la connaissance de la valeur phonémique (Ecalle, 2004). Lorsque le phonème est en position initiale dans le nom de la lettre (b /*bi*/), sa valeur phonémique est mieux retenue que lorsqu'il est en position finale (f /ɛf/). D'autres résultats suggèrent que des compétences phonémiques interviendraient également dans le processus d'extraction du son des lettres à partir de leur nom. Foulin (1995) montre que la connaissance du nom des lettres accélère le développement de la connaissance du son des lettres et la conscience phonémique qui à leur tour sous-tendraient l'acquisition du principe alphabétique de la lecture. L'évolution de la connaissance du nom des lettres pourraient favoriser le développement simultané et ultérieur de la sensibilité phonémique, ce qui expliquerait l'aspect prédicteur du niveau de connaissance du nom des lettres sur les performances phonémiques pendant la période préscolaire et les premières années de scolarité (Foulin, 2007).

L'étude de Foulin & Pacton (2006) confirme que la connaissance du nom des lettres peut jouer un rôle actif dans l'apprentissage initial de l'écrit et qu'elle n'est pas seulement un outil de dénomination et de manipulation des lettres. Les résultats ont montré que les enfants pré-lecteurs âgés de 5-6 ans étaient capables d'utiliser des relations lettres-phonèmes pour distinguer des mots écrits. Avant l'apprentissage formel de la lecture, la dénomination de lettres est un prédicteur de la lecture plus puissant que la connaissance du son des lettres (Schatschneider & al., 2004). Muter & al. (1998) montrent qu'une part de variance des performances en lecture en 1ère année est expliquée par la connaissance des lettres puis par la conscience phonémique. En 2ème année, les performances en lecture sont expliquées par celles obtenues l'année précédente et également par la connaissance des lettres au même moment.

La recherche de Biot-Chevrier, Ecalle & Magnan (2008) se proposait d'étudier l'effet de la connaissance du nom des lettres dans le traitement précoce des relations oral-écrit auprès

d'enfants francophones âgés de 4 ans scolarisés en moyenne section de maternelle. Les résultats révèlent que la connaissance des lettres participe aux premières productions orthographiques. D'une part, les enfants *connaisseurs* semblent saisir que les relations entre l'oral et l'écrit s'effectuent sur la base des lettres et d'autre part, le répertoire de lettres qu'ils ont pu constituer leur permet d'y sélectionner la lettre correspondant le mieux à la séquence orale qu'ils perçoivent et identifient dans le mot prononcé. Les enfants *connaisseurs* manifestent également des compétences phonologiques plus développées.

Dans une étude longitudinale, Ecalle, Magnan & Biot-Chevrier (2008) ont examiné la contribution de la connaissance du nom des lettres (CNL) et de la conscience phonologique mesurés en maternelle sur les tâches de lecture et d'écriture en CP. Les résultats montrent que les enfants ayant les meilleures performances dans la CNL en maternelle sont ceux qui ont obtenus les meilleurs résultats dans les épreuves de lecture de mots et de pseudo-mots et l'épreuve d'écriture présentées en 1ère année. Les analyses de régression révèlent que la connaissance du nom des lettres explique 6% de variance des performances de lecture de pseudo-mots et 4% des performances en identification de mots. Cette étude distingue entre deux aspects de connaissance du nom des lettres : un aspect prédictif et un aspect évaluatif. Le premier indique que le poids prédictif de la CNL sur le niveau ultérieur en lecture et écriture est faible alors que le second permet de considérer la CNL comme un précurseur de la connaissance du code alphabétique.

La recherche de Bonnefoy & Rey (2008) souligne l'importance d'une bonne connaissance du nom des lettres au tout début de l'apprentissage de la lecture. Ils ont mesuré chez 54 enfants français de CP d'âge moyen 6 ans 6 mois leur connaissance du nom des lettres de l'alphabet. Les enfants ont tous réalisé une épreuve de connaissance de lettres (ils doivent dénommer à voix haute la lettre qui est présentée à l'écran), une épreuve de lecture de mots (ils doivent prononcer à voix haute le mot qui apparait à l'écran le plus correctement et le plus rapidement possible), un test de vocabulaire et un test de rapidité pour estimer leur vitesse de traitement de l'information. Les résultats indiquent que la connaissance des lettres explique de manière indépendante une part importante de variance des performances en lecture. Si bien connaître le nom des lettres semble être un facteur facilitateur pour l'acquisition de la lecture, les auteurs remarquent aussi qu'automatiser cette connaissance semble tout aussi déterminant. La connaissance des lettres mais aussi la capacité à les dénommer rapidement et automatiquement semblent donc deux facteurs importants pour le développement de la lecture.

Dans l'ensemble, ces résultats pourraient suggérer des interventions précoces auprès d'enfants à risque de difficulté en lecture pour faciliter l'apprentissage des lettres, premier pas vers la compréhension du principe alphabétique.

4. La dénomination rapide

Les épreuves de dénomination rapide ou RAN (pour Rapid Automatized Naming) ont pour objectif de tester la précision et la rapidité de l'accès au lexique d'items très fréquents. Les items peuvent être des chiffres, des lettres, des couleurs ou des images. Cette mesure fournit ainsi une estimation de la capacité d'un enfant à retrouver et à produire rapidement le nom d'un item à partir du symbole visuel qui le représente. Autrement dit, s'il dispose d'une connaissance très stable et automatisée de cette association et également, s'il est en mesure de la produire rapidement (Bonnefoy & Rey, 2008).

Les études longitudinales montrent que les performances en dénomination rapide mesurés chez les prélecteurs sont prédictives de leurs performances ultérieures en lecture (Kirby, Parilla & Pfieffer, 2004 ; Lecoq, 1991 ; Piquard-Kipffer, 2003 ; Scarborough, 1998). Badian (1993 ; cité par Castel & al. 2008) a montré que la dénomination rapide de lettres est un bon prédicteur de la reconnaissance de mots tandis que la dénomination rapide d'objets est un bon prédicteur de la compréhension de texte. D'autres études ont montré un déficit de la dénomination rapide chez les sujets dyslexiques (Castel, Pech-Georgel, George & Ziegler, 2008). Selon la théorie du double déficit de Wolf & Bowers (1999), les enfants dyslexiques peuvent être subdivisés en trois groupes distincts selon leurs performances dans les tâches de conscience phonologique et de RAN : ceux qui présentent un déficit dans la tâche de conscience phonologique, ceux qui présentent un déficit dans la tâche de RAN et ceux qui présentent un déficit dans la tâche de conscience phonologique et de RAN.

Même si la relation entre le RAN et la lecture a été démontrée dans de nombreuses études, il reste néanmoins à identifier les mécanismes cognitifs sous-jacents au RAN qui sont également impliqués dans l'acquisition de la lecture. Certains auteurs considèrent, en effet, que le RAN et la lecture partagent certains processus visuo-attentionnels. La dénomination d'un élément parmi d'autres met en jeu des processus de balayage visuel, de contrôle oculaire ou encore d'attention focalisée. Tous ces processus sont également impliqués lorsqu'un enfant lit un texte (Bosse, valdois & Dompnier, 2009). Une autre explication pour l'existence d'un lien entre RAN et lecture repose sur l'idée que la dénomination rapide met essentiellement en

jeu des traitements phonologiques : le RAN mesurerait l'accès aux représentations phonologiques.

Dans le but de mieux comprendre le lien entre RAN et lecture, Castel & al. (2008) ont réalisé une expérience auprès d'enfants dyslexiques dans laquelle ils se sont intéressés à l'importance des facteurs visuo-attentionnels et des facteurs d'automatisation dans le paradigme du RAN. Dans un premier temps, ils ont comparé les performances d'enfants dyslexiques (N=24) et normolecteurs (N=24) dans une version continue et une version discrète. Dans la version continue, tous les items à dénommer sont présentés au même temps sur papier (version classique), tandis que dans la version discrète, les items sont présentés successivement réduisant ainsi l'implication des facteurs visuo-attentionnels. Les résultats suggèrent que la tâche de RAN permettait de discriminer les enfants dyslexiques des enfants normolecteurs. En effet, 62% des enfants dyslexiques présentent un déficit en dénomination rapide. Ce déficit est identique quel que soit la version de la tâche excepté pour les chiffres où le déficit est plus important dans la version discrète que dans la version continue du RAN. Les facteurs visuo-attentionnels (contrôle des mouvements oculaires ou attention focalisée) semblent donc ne pas être à la base de ce déficit. Les résultats montrent également que la dénomination rapide de stimuli très automatisés tels que les chiffres ou les lettres est plus facile que la dénomination rapide de stimuli moins automatisés comme les objets. En effet, moins les items sont automatisés, plus les informations non pertinentes peuvent interférer avec le processus de dénomination. En revanche, la comparaison des différents types de stimuli ne permet pas d'expliquer à eux seuls le déficit obtenu dans le RAN. Cette recherche montre que la force du RAN comme prédicteur de la lecture réside avant tout dans sa capacité à mesurer l'accès rapide aux représentations phonologiques. Les analyses de régression montrent que chacune des deux variables (RAN et conscience phonologique) prédit une part unique de variance des performances en lecture. RAN et lecture partagent donc un mécanisme fondamental : la mise en relation de l'automatisation du couplage entre les représentations visuelles et phonologiques.

5. Compétences cognitives et apprentissage de la lecture

Comprendre les difficultés d'apprentissage de la lecture requiert en outre de s'intéresser aux processus cognitifs associés à cet apprentissage. Parmi ces processus, il convient de souligner l'importance des processus mnésiques, des fonctions excécutives, du niveau intellectuel et des capacités visuo-attentionnelles.

5.1. La mémoire de travail

Bien que la majorité des études ait traité de la relation entre lecture et habileté métaphonologiques, d'autres types de traitements phonologiques sont associés à la lecture. Plusieurs recherche affirment l'existence d'une relation entre le développement de la lecture et le codage phonologique en mémoire de travail, c'est-à-dire le fait de coder l'information sous forme de parole interne pour pouvoir la stocker efficacement (voir Jorm, 1983). La mémoire de travail est réputée être impliquée dans l'acte de lire tant au niveau de l'identification des mots qu'au niveau des processus de compréhension (Demont, 2002). Cette relation s'expliquerait par le fait qu'un codage efficace en mémoire de travail permettrait au lecteur débutant de maintenir une représentation exacte de phonèmes associés aux lettres ou groupes de lettres. Plus particulièrement, l'enfant doit maintenir en mémoire les lettres et fragments de mots déjà identifiés pendant qu'il décode simultanément les lettres restantes (Wagner & Torgesen, 1987).

La mémoire de travail[26] est définie comme « *un système de maintien temporaire et de manipulation de l'information nécessaire pour réaliser des activités cognitives complexes* » (Baddley, 1986 ; p. 79). Selon le modèle développé par Baddley & Hitch (1974), la mémoire de travail se compose d'un processeur central et de deux sous-systèmes périphériques (visuo-spatial et verbal). Le processeur central est responsable essentiellement du traitement de l'information et la coordination des systèmes périphériques. Les deux sous-systèmes périphériques servent à stocker et à maintenir temporairement l'information. Ils assurent soit le stockage des informations visuelles et spatiales sous forme d'images visuo-spatiales, soit le stockage des informations verbales sous forme phonologique. L'information verbale serait stockée selon ses caractéristiques phonologiques au niveau de la boucle articulatoire et maintenue active grâce au processus d'autorépétition mentale. Les différents travaux suggèrent que les déficits au niveau de la mémoire de travail peuvent expliquer en partie les difficultés en lecture. D'ailleurs, la performance à des tâches impliquant la récupération immédiate et ordonnée des stimuli verbaux au niveau préscolaire est un prédicteur des différences individuelles en lecture à la fin de la première année scolaire (Demont, 1992 ; Lecoq, 1991 ; Mann & Liberman, 1984). La mémoire phonologique est susceptible de contribuer au développement de la procédure phonologique de deux manières ; en permettant

[26] On distingue la mémoire à court terme ou mémoire immédiate, symbolisant un stockage temporaire d'éléments récemment encodés maintenus par des stratégies simples comme la répétition et la mémoire de travail, faisant appel aux processus attentionnels exigés par un traitement et une manipulation de l'information dans le but de stocker une partie ou le produit final de ces traitements (Baddley, 1986).

d'une part le stockage temporaire de l'information phonologique lors du décodage et en jouant d'autre part un rôle dans l'apprentissage à long terme des règles de CGP indispensables à son développement (Gathercole & Baddley, 1993). La mémoire phonologique est normalement évaluée par des tâches qui demandent de retenir pour un bref moment des séquences d'items verbaux avec ou sans signification. L'explication la plus communément acceptée des difficultés dans ce type de tâche (empan de mémoire et mémoire de travail) implique des problèmes de représentation mentale des aspects phonologiques du langage.

Grâce à une comparaison transversale, Siegel & Ryan (1989) tentent de valider l'existence de différences individuelles en mémoire de travail chez des enfants jugés en avance ou en retard dans la reconnaissance des mots. Les auteurs utilisent deux indicateurs de l'empan en mémoire de travail dont le premier est linguistique et le second est non linguistique (rappel dans l'ordre de figures géométriques). Les différences individuelles observées dans les deux épreuves et aux trois âges (7-8 ans, 9-10 ans et 11-13 ans) indiquent que les mauvais lecteurs ont les performances les plus faibles. Ces résultats laissent supposer que c'est le processeur central qui ne remplirait pas efficacement son rôle.

Afin d'examiner les liens entre conscience phonologique, mémoire de travail et l'acquisition de la lecture, Rohl & Pratt (1995) ont réalisé une étude longitudinale avec 76 enfants. L'expérimentation s'est déroulée en trois temps. Au début de la première année (temps 1), trois mesures de mémoire (mémoire de lettres, mots et phrases) et deux mesures des compétences phonologiques (catégorisation et segmentation phonémique) ont été administrées aux enfants (âge moyen 5 ans 8 mois). Au temps 2, c'est-à-dire à la fin de la 1ère année, les enfants ont repassé les différents tests de mémoire et de phonologie. Ils ont reçu également, de façon individuelle, un test de fusion de phonèmes, de lecture de mots et de pseudo-mots et un test de compréhension. La dernière évaluation est effectuée au second semestre de la 2ème année (temps 3). Les tests de mémoire et de lecture sont réadministrés exactement comme à la première et à la seconde phase. Les analyses de régressions multiples indiquent que les compétences phonologiques précoces prédisent le niveau ultérieur en lecture même lorsque la capacité en mémoire de travail est contrôlée. En revanche, la mémoire de travail ne contribue pas à l'acquisition de la lecture en 2ème année indépendamment des compétences phonologiques. Par ailleurs, la mémoire du travail, et particulièrement la répétition de mots à l'envers, mesurée en 1ère année prédit les performances en lecture en 2ème année. Toutefois, lorsque les trois variables de compétences phonologiques sont contrôlées, la répétition de mots à l'envers ne contribue pas de façon significative à l'acquisition de la

lecture en 2ème année. Les résultats de cette étude confirment la relation étroite entre conscience phonologique et mémoire de travail, notamment au début de l'apprentissage de la lecture. Alors que les compétences phonologiques mesurées au début de la 1ère année apportent une contribution significative à l'acquisition de la lecture à la fin de la 1ère année, la mémoire de travail semble prédire le niveau de lecture à des stades plus tardifs.

Demont (2003) s'est intéressée à l'étude simultanée des processus de traitement des unités phonologiques et des processus de stockage et de traitement en mémoire de travail chez des enfants présentant des difficultés en lecture. L'étude a porté sur vingt-deux enfants dyslexiques d'âge moyen 10 ans 3 mois présentant de graves difficultés dans le langage écrit et quarante-huit enfants scolarisés en première année d'apprentissage de la lecture d'âge moyen 6 ans 9 mois. En fonction de leur performance à une épreuve de lecture, ces derniers sont répartis en trois sous-groupes de lecteurs : faibles lecteurs, normolecteurs et bons lecteurs. Outre les épreuves de lecture et de conscience phonologique, différentes épreuves mnésiques évaluant les fonctions de stockage et de traitement de la mémoire de travail sont présentées aux enfants : empan de mots, empan de chiffres à rebours et empan de phrases. Les résultats indiquent que les bons lecteurs obtiennent des performances significativement supérieures à celles des enfants dyslexiques, des normolecteurs et des faibles lecteurs. De plus, les scores des bons lecteurs ne témoignent pas d'un effet de similarité phonologique (empan similaire quelle que soit la nature des mots). Par contre, les enfants dyslexiques apparaissent, quant à eux être sensibles à la similarité phonologique des items à rappeler (empan de mots phonologiquement dissemblables supérieur à l'empan de mots phonologiquement semblables). Ces données semblent ainsi révéler un dysfonctionnement au niveau de la boucle articulatoire et non d'un déficit au niveau de la capacité de stockage. Bien que les différentes mesures mnésiques apparaissent corrélées avec les capacités de reconnaissance des mots, seul l'empan de mots phonologiquement similaires contribue significativement à ces capacités. Néanmoins, sa contribution n'apparaît pas indépendante de celle des mesures de conscience phonologique. Par ailleurs, les processus mnésiques responsables du traitement et de la gestion de l'information verbale (évaluée à l'aide de l'épreuve de l'empan de phrases) contribuent quant à eux de façon significative et indépendante aux performances en compréhension.

Lecoq (1991) a proposé à près de 300 enfants de 42 à 78 mois douze épreuves mnésiques. Il visait à recueillir des informations sur l'empan mnésique en fonction de certaines caractéristiques des items, en l'occurrence, leur longueur, leur similarité phonétique ou

sémantique, leur familiarité et leur catégorie (chiffres ou mots). A la différence des épreuves d'analyse phonologique, il observe peu de progrès pour les capacités de mémoire entre 3 ans et demi et 6 ans. De plus, les disparités entre les épreuves pour les différentes tranches d'âge sont minimes. Les épreuves les mieux réussies sont celles qui impliquent le rappel des phrases rimantes et non rimantes et le rappel de chiffres à l'envers. Par ailleurs, aucune épreuve ne différencie nettement les bons lecteurs des lecteurs moins compétents de la maternelle au cours élémentaire. Selon Lecoq, il semble que l'apprentissage de la lecture accroissent les capacités mnésiques des sujets en ce qui concerne au moins le matériel verbal.

5.2. Rôle des fonctions exécutives

Comme toutes les activités complexes, la lecture exige une régulation (Gombert & Fayol, 1995). Cette régulation est importante chez l'apprenti lecteur car le début de l'apprentissage est particulièrement complexe. Une hypothèse récente accorde un rôle primordial au contrôle exécutif (CE) dans l'apprentissage de la lecture (Boulc'h, 2008 ; Boulc'h, Gaux & Boujon, 2007). Le CE « *se réfère à plusieurs habiletés qui permettent d'assurer le bon déroulement d'une activité en initiant l'action et en sélectionnant la stratégie la plus appropriée et particulièrement dans les situations complexes où des réponses automatiques sur-apprises ou des habitudes peuvent entraver son déroulement* » (Boulc'h, 2008, p. 81). Il intervient dans les tâches cognitivement coûteuses qui nécessitent un traitement complexe ou qui comportent une interférence entre des stimuli ou des réponses en compétition notamment l'alternance, l'inhibition et la mise à jour en mémoire du travail.

De nombreux travaux montrent l'implication du CE dans les tâches complexes de mémorisation (Swanson, 1999). Cependant, peu de travaux ont jusqu'alors étudié son rôle lors de la lecture. Le CE pourrait être impliqué dans l'identification des mots en permettant d'alterner d'une part les traitements de conversion grapho-phonémiques et d'assemblage des unités sonores converties et d'autre part, les différentes procédures de décodage (analyse des graphèmes ou des syllabes). Ce rôle pourrait être particulièrement critique à certaines étapes de la lecture pour les faibles lecteurs qui n'ont pas encore totalement automatisé leurs procédures d'identification de mots et qui mobilisent fortement de ce fait la composante exécutive de la mémoire de travail (Boulc'h, Gaux & Boujon, 2007).

L'implication des fonctions exécutive dans l'activité de lecture a été étudié par Boulc'h, Gaux & Boujon (2007). Pour se faire, ils ont demandé à des normolecteurs et faibles lecteurs de CE2 d'effectuer une tâche de conjonction illusoire (CI) qui nécessite la sélection d'une

caractéristique d'une lettre appartenant à un graphème ou se situant à la frontière syllabique d'un mot. Cette tâche est réalisée seule ou en même temps qu'une seconde tâche exécutive de mise à jour ou d'alternance. Partant du postulat que les processus exécutifs se manifesteraient davantage lorsque le traitement cognitif est plus coûteux, deux versions de la tâche de CI ont été utilisées, l'une induisant l'utilisation des graphèmes et l'autre induisant l'utilisation des syllabes. Les résultats confirment l'implication du contrôle exécutif : l'ajout d'une tâche de contrôle exécutif augmente le nombre d'erreurs et les temps de réponse chez les faibles lecteurs de CE1. Les faibles lecteurs sont systématiquement en difficulté lorsqu'un contrôle (alternance, mise à jour, inhibition) est exigé. Ils auraient ainsi des difficultés à contrôler les processus mis en jeu dans une tâche complexe comme la lecture coûteuse pour les faibles lecteurs. Ce déficit de contrôle s'additionnerait à de faibles capacités de traitement phonologique. Le coût engendré par l'ajout d'une tâche exécutive apparaît davantage sur les temps de réponse. Les analyses des données comparant chacune des conditions avec contrôle exécutif à la condition contrôle de CI indiquent que le coût du contrôle exécutif n'est pas significatif pour les normolecteurs, alors que ce coût se manifeste chez les faibles lecteurs.

5.3. Le niveau intellectuel

Le lien entre intelligence et apprentissage de la lecture est souvent l'objet de controverses. Les mesures d'intelligence ne permettent d'expliquer qu'une très faible partie des différences individuelles en capacité de lecture au cours préparatoire. Un certain nombre d'études ont montré qu'il n'existe apparemment pas de relation, au moins directe, entre le niveau d'intelligence des enfants et leur niveau de réussite en lecture. Un niveau faible de QI ne s'accompagne pas systématiquement d'un mauvais niveau en lecture (Siegel, 1988, 1989 ; Rodrigo & Jimenez, 2000). Ceci est sans doute à imputer à la définition classique de la dyslexie développementale qui présente les dyslexiques comme « *une minorité d'enfants d'intelligence normale et en bonne santé mais qui échouent dans l'apprentissage de la lecture* » (Sprenger-Charolles & Colé, 2003 ; p. 22). Comme le souligne Siegel (1990), l'intelligence ne peut être mesurée indépendamment de la réussite scolaire. Les QI obtenus à partir de tests standardisés mesurent l'efficacité de fonctions telles que la mémoire de travail ou le langage. De fait, un bas niveau de QI peut être la conséquence de ces difficultés et non pas une mesure précise du niveau d'intelligence de l'enfant.

Lecoq & al. (1996) ont effectué une recherche corrélationnelle ayant pour but d'étudier les liens susceptibles d'exister entre la phonologie, la mémoire, l'intelligence et la lecture chez

des enfants de différents niveaux de scolarité. Les résultats montrent que les performances en lecture ne sont pas très fortement liées au niveau intellectuel au début de l'apprentissage mais les corrélations augmentent sensiblement en fonction du niveau scolaire. Cette augmentation peut s'expliquer, en grande partie par la relation causale inverse. En lisant mieux et en lisant plus, on développe ses bases de connaissances, ses procédures de traitement de l'information, et on devient plus performants dans les tests dits d'intelligence (Observatoire National de la lecture, 1998).

Jimenez, Siegel & Lopez (2003) ont tenté d'éclaircir les relations entre QI et apprentissage de la lecture avec des enfants espagnols en difficulté en fonction de leurs scores de QI verbal, performance et global. Les auteurs observent des performances en lecture similaires quel que soit le niveau de QI. Le concept même d'intelligence reste controversé et il n'existe pas d'accord sur sa nature (Stanovich, 1989). Perfetti (1985) considère que l'aptitude à comprendre ce qu'on lit ne se distingue pas de l'aptitude intellectuelle générale surtout lorsque celle-ci est évaluée sous forme verbale. L'intelligence est assimilée à l'étendue des connaissances et des expériences dont le lecteur dispose. L'auteur explique ainsi les corrélations observées entre tests d'intelligence et tests de lecture.

Les recherches actuelles soutiennent l'idée selon laquelle les difficultés d'apprentissage de la lecture ne proviendront pas de problèmes intellectuels mais d'une déficience des traitements phonologiques du langage. Ainsi, Stanovich, Cunningham & Cramer (1984) constatent que les tâches phonologiques mesurées chez des enfants prélecteurs sont significativement corrélées avec le niveau de lecture atteint en première année et qu'elles constituent également de meilleurs prédicteurs de la réussite en lecture que le QI global puisqu'elles contribuent à plus de 66% de la variance en lecture. Dans le même sens, les résultats de l'étude de Siegel (1993), portant sur un échantillon de 1493 enfants âgés de 7 à 16 ans normolecteurs et en difficulté de lecture, montrent que la capacité à lire les pseudo-mots constitue le facteur ayant la part explicative la plus importante de la reconnaissance des mots écrits et de la compréhension, alors que le QI n'explique qu'une part très réduite de cette variance. Les traitements phonologiques semblent donc être à l'origine des difficultés en lecture et leur mise en œuvre serait relativement indépendante de celle de processus cognitifs plus généraux. Scarborough (1989) a suivi longitudinalement 66 enfants de l'âge de 5 ans jusqu'à l'âge de 8 ans. Parmi ces enfants, 38 sont issus de familles « à risque de dyslexie » et 28 enfants sont issus de « familles contrôles ». Les résultats indiquent qu'en 2ème année, le groupe d'enfants ayant développé une dyslexie manifeste des niveaux de QI significativement inférieurs à ceux

du groupe contrôle. Cependant, cette différence est due au QI verbal des enfants dyslexiques. Ainsi, il semble qu'un accès limité à la stimulation langagière pourrait, dans certains cas, infléchir le développement verbal.

5.4. Les capacités visuo-attentionnelles

Le rôle spécifique de l'attention visuelle dans l'acquisition de la lecture est peu étudié. Certaines études montrent que la variabilité observée au niveau du traitement phonologique ne peut pas expliquer l'ensemble des variabilités observées dans les capacités de lecture des enfants et suggèrent que d'autres processus cognitifs sont également essentiels pour l'acquisition de la lecture (Valdois, 2003 ; Bosse, Valdois & Dompnier, 2009). Valdois (2005) considère qu'une distribution non adéquate de l'attention engendrait des difficultés de lecture car certaines lettres apparaîtraient comme plus saillantes (par exemple : MAisoN). Lire consiste à porter son attention sur les différents mots et à les sélectionner individuellement l'un après l'autre pour pouvoir les identifier correctement. Les traitements s'exercent ensuite sur les différentes lettres constituant le mot pour qu'elles puissent être traitées en même temps. La notion d'empan visuo-attentionnel (EVA) renvoie à la quantité d'éléments visuels distincts pouvant être traités simultanément ; elle correspond en lecture au nombre d'unités orthographiques qui peuvent être analysées en parallèle au sein de la séquence du mot (Colé & Valdois, 2007).

L'hypothèse visuo-attentionnelle soutient que la création en mémoire d'une trace orthographique du mot n'est possible que si l'enfant est capable de distribuer son attention visuelle de façon harmonieuse sur l'ensemble des lettres qui composent le mot (Valdois, 2003). En situation de lecture de texte, l'attention doit se porter sélectivement et successivement sur chacun des mots individuels pour que les procédures d'identification de mots puissent s'appliquer. Au niveau du mot lui-même, des traitements visuo-attentionnels entrent en jeu de façon à traiter l'ensemble des lettres de la séquence. Il est notamment nécessaire que l'enfant distribue son attention visuelle de façon équitable sur l'ensemble des lettres qui composent le mot, pour que les lettres puissent être traitées en parallèle et le mot correctement identifié. Cette hypothèse est actuellement appuyée par des études avec des enfants dyslexiques qui montrent que les troubles d'acquisition de la lecture peuvent exister conjointement à des déficits visuo-attentionnels (Valdois, 2003 ; Valdois, Bosse & Tainturier, 2004). Plusieurs études récentes ont suggéré que le traitement visuo-attentionnel effectué sur les mots vus pouvait avoir une influence sur la lecture et l'orthographe des enfants

dyslexiques (Bosse & Valdois, 2007). Bosse, Tainturier & Valdois (2007) ont montré qu'un déficit de l'EVA contribuait à la faible performance des enfants dyslexiques indépendamment de leurs aptitudes phonologiques.

Les données actuelles permettent donc de penser qu'un déficit visuo-attentionnel constitue une origine plausible des troubles dyslexiques, distincte de l'origine phonologique. L'étude spécifique d'enfants dyslexiques, présentant un retard important d'apprentissage de la lecture en dépit de bonnes aptitudes métaphonologiques et langagières, a conduit certains auteurs à faire l'hypothèse de troubles visuo-attentionnels associés (Bosse & al., 2007 ; Valdois & al., 2003 ; Valdois, Bosse & Tainturier, 2004). Les études menées depuis ont permis de montrer que ces troubles sont relativement fréquents dans la population dyslexique et qu'ils se rencontrent souvent indépendamment de toute atteinte phonologique. L'étude de Bosse & al. (2007) montre une forte corrélation entre les performances des enfants dyslexiques sur les épreuves visuo-attentionnelles et leurs performances dans la lecture de mots isolés. Des analyses de régressions multiples ont par ailleurs montré que leurs capacités de traitement visuo-attentionnel étaient prédictives de leur niveau de lecture indépendamment de leurs capacités de traitement phonologique. Ces résultats suggèrent qu'un déficit de l'empan visuo-attentionnel explique de façon spécifique le faible niveau de lecture des enfants dyslexiques.

Bosse & Valdois (2009) ont conduit une étude longitudinale qui a porté sur 417 élèves évalués en 1ère, en 3ème et en 5ème année dans le but d'examiner le rôle de l'EVA dans le développement des performances en lecture. Une batterie incluant des tâches de lecture, de conscience phonémique et d'EVA a été administrée aux enfants. Les résultats montrent que l'EVA prédit les performances en lecture indépendamment de la conscience phonémique. En revanche, alors que l'influence spécifique de l'EVA sur la lecture de pseudo-mots diminue de la 1ère année jusqu'à la 3ème année, cette tâche continue à contribuer à la lecture des mots irréguliers à tous les niveaux scolaires. Ainsi, tout comme la conscience phonologique, l'EVA contribue de façon indépendante à l'explication de la variance en lecture dès le début de l'apprentissage et semble avoir un effet à long terme sur l'acquisition des connaissances orthographiques.

Bosse, Valdois & Dompnier (2009) ont essayé de confirmer l'existence d'un lien causal entre empan visuo-attentionnel et acquisition de la lecture en étudiant la capacité prédictive du traitement visuo-attentionnel mesuré précocement sur l'acquisition du langage écrit un an plus tard, tout en prenant en compte les autres facteurs prédictifs connus, essentiellement les

traitements phonologiques. L'étude a porté sur 127 enfants de grande section de maternelle (GSM) d'âge moyen 6 ans 10 mois. Tous les enfants ont été testés individuellement trois fois : en GSM, en CP et en CE1. Quatre types de tâches ont été proposés : des tâches de lecture, des épreuves visuo-attentionnelles et des tâches contrôle. Les analyse de corrélation montrent que les deux mesures de lecture, en CP et CE1, sont corrélées aussi bien au score phonologique (r=.51 et r=.38) qu'au score visuo-attentionnel (r=.56 et r=.43). Les analyse de régression multiple montrent que l'empan visuo-attentionnel, mesuré en CP, prédit effectivement une part propre de variance en lecture (75%) un an après. Cette part est indépendante de la part de variance expliqué par la conscience phonologique en CP. Ces résultats sont compatibles avec l'hypothèse selon laquelle l'empan visuo-attentionnel mesure une capacité cognitive précoce directement reliée à l'acquisition ultérieure de la lecture, indépendamment du lien existant entre traitement phonologique et lecture. L'hypothèse dépasse cependant les seuls cas de dyslexie développementale. Il est probable que, comme le traitement phonologique, le traitement visuo-attentionnel soit variable d'un lecteur à l'autre. Dans ce cas, on s'attend à ce que de bonnes capacités de traitement visuo-attentionnel s'accompagnent de bonnes capacités d'acquisition des connaissances orthographiques. Certaines études suggèrent effectivement un lien entre traitement visuo-attentionnel et lecture. Casco, Tressoldi, & Dellantonio (1998, cité par Bosse, 2005) montrent que les enfants les plus lents dans une tâche de barrage de cibles sont également plus lents que les autres en lecture et font un nombre d'erreurs visuelles plus élevé. Ainsi, la performance dans une tâche impliquant l'attention visuelle sélective semble liée à la performance en lecture.

Globalement, ces résultats suggèrent que les troubles visuo-attentionnels sont assez fréquemment associés aux troubles dyslexiques et que ces troubles sont le plus souvent observés indépendamment des troubles phonologiques. Si certains processus de traitement et certains types de représentations ont une importance déterminante dans l'appropriation de l'écrit par l'enfant, comment les méthodes pédagogiques de l'enseignement de la lecture et les caractéristiques linguistiques de chaque langue influencent-elles le développement de ces mécanismes et de ces connaissances ?

6. Influence des méthodes d'enseignement

L'apprentissage de l'écrit ne se réalise pas de manière spontanée et naturelle comme l'apprentissage du langage oral (Gombert, 1992). En particulier, la découverte du code alphabétique nécessite le plus souvent l'intervention explicite d'une instruction formelle.

Etant donné la nature des relations entre la chaine écrite et la parole, on peut concevoir deux manières d'obtenir de l'information sur la prononciation d'un mot écrit. L'une, souvent nommée adressage qui consiste à rechercher un appariement direct entre le mot écrit et sa représentation orthographique dans le lexique mental. La seconde procédure, appelée assemblage, consiste à tirer partie des régularités présentes dans tout système d'écriture phonographique entre les éléments de la chaine écrite et de la chaine parlée. Certains chercheurs (Baron & Treinman, 1980) ont avancé l'idée selon laquelle, à niveau d'efficience égal, certains lecteurs utiliseraient principalement l'assemblage et d'autres, principalement l'adressage. L'examen de l'influence éventuelle des méthodes d'enseignement est susceptible d'apporter une explication. Snowling (1996) suggère que la majorité des enfants apprendrait à lire avec succès quel que soit l'enseignement dispensé mais, à l'inverse, que le choix de la méthode d'apprentissage serait crucial pour les enfants à risque d'échec en lecture ou qui manifestent des difficultés en lecture. Le problème récurrent qui se pose concerne le type de méthode d'enseignement proposé aux enfants dans le cadre de l'apprentissage institutionnelle de la langue écrite et son efficacité.

Avec l'introduction de l'école obligatoire en Tunisie, diverses méthodes pédagogiques ont été proposées pour l'enseignement de la lecture : synthétique[27], ensuite globale et plus récemment mixte. Actuellement, c'est la méthode mixte qui est adopté dans l'apprentissage de la lecture. Elle s'articule autour de trois étapes : une première étape de préparation à la lecture à l'aide de la lecture d'unités globales (phrases, énoncés, mots) et en se basant sur des moyens variés tels que les images, le contexte et la comparaison entre les connaissances déjà acquises et celle que l'élève va acquérir. Au cours de la deuxième étape qui vise le déchiffrage, la lecture se fait par une imprégnation visuelle sur un matériel global. Ensuite, on descend aux phrases, aux mots puis aux syllabes et enfin aux lettres. Dans la troisième étape, la lecture est conçue dans ses aspects les plus variés. C'est une opération de déchiffrage et aussi un moyen de compréhension et d'accès à l'information. Le but de l'enseignement de l'arabe est de donner aux élèves la capacité de comprendre, de s'exprimer et, éventuellement, de penser dans cette langue[28]. La tâche est rendue complexe par la coexistence à l'intérieur de la langue arabe de deux registres : l'arabe dialectal et l'arabe standard. L'arabe standard est réservé généralement à l'écriture et enseigné prioritairement à l'école mais les élèves ont d'autant plus facilement accès à la communication orale, celle du dialogue familier et quotidien c'est à dire à l'arabe

[27] L'objectif de cette méthode est d'entrainer l'élève à la prononciation exacte et à la bonne récitation
[28] Ministère de l'Education Nationale. *Les programmes officiels pour l'enseignement de la langue arabe* (œuvre publiée en arabe).Tunis, 2000.

dialectal. C'est la raison pour laquelle les méthodes actuelles laissent apparaître une préoccupation de plus en plus accentuée du problème de la diglossie en essayant de trouver un compromis entre le langage de l'enfant en arabe dialectal et le langage enseigné à l'école en arabe standard. Les méthodes proposées ont tenté, d'une part, d'étayer l'apprentissage de la lecture sur les acquis linguistiques de l'enfant et de rétrécir les distances entre l'arabe dialectal et l'arabe standard, d'autre part.

Content & Leybeart (1992) ont essayé d'établir dans quelle mesure les caractéristiques des méthodes d'enseignement de la lecture peuvent influencer la nature des mécanismes mis en œuvre dans la reconnaissance des mots écrits. Cette recherche avait pour objectif d'évaluer les effets de deux grandes méthodes d'enseignement appelées méthodes globales et méthodes phoniques ou analytiques. Les méthodes phoniques se caractérisent par le fait qu'on enseigne progressivement, dès le début de la première année, des correspondances entre lettres et sons du langage. Dans les méthodes globales, le début de l'apprentissage est basé sur la mémorisation des formes écrites des mots. L'enfant identifie les mots à partir de leur forme globale tout en accédant au sens du texte[29]. Les correspondances lettres-sons ne sont pas enseignées mais introduites lorsque l'occasion se présente, et l'accent est mis sur la compréhension et la communication. L'hypothèse de cette recherche est que les méthodes phoniques favorisent le développement précoce de stratégies basées sur l'assemblage ; en revanche, les méthodes globales devraient favoriser le développement du lexique logographique et l'utilisation de processus d'dressage. Plus précisément, la méthode phonique favoriserait le développement de la procédure phonologique et la méthode globale celui de la procédure orthographique. L'étude a porté sur des enfants de trois niveaux scolaires (2ème, 4ème et 6ème année) dont les uns avaient reçus un enseignement selon la méthode phonique et les autres, selon la méthode globale. Les résultats obtenus tendent à montrer, de manière générale, que quelle que soit la méthode d'apprentissage de la lecture utilisée au début du cursus scolaire, on observe une amélioration de la rapidité et de l'efficience de l'assemblage phonologique et un développement progressif du recours à l'adressage orthographique direct. Les auteurs constatent, en effet, que non seulement tous les enfants ont recours progressivement à la procédure par médiation phonologique mais que les enfants assignés à la méthode globale présentent des performances, en 2ème année, qui indiquent à la fois un

[29] Le ministre français de l'Education Nationale a décidé de proscrire, dès la rentrée scolaire 2006, la méthode globale d'apprentissage de la lecture et de la remplacer par la méthode dite syllabique. Il s'agit, selon lui, d'une source de difficulté chez de nombreux élèves et de l'une des causes fondamentales de l'échec dans l'apprentissage de la lecture.

recours massif à la procédure phonologique et un recours extrêmement limité à la procédure orthographique.

Le travail de Goigoux (2000) tente d'apporter des éléments de réponses plus tranchés sur l'effet des méthodes globales et phoniques. Pour lui, ce qui distingue ces deux méthodes, ce n'est pas la présence ou l'absence d'un enseignement explicite des CGP mais leur démarches d'analyse du langage qui, pour les méthodes globales, débute avec les unités les plus larges (phrases, mots) pour parvenir aux plus élémentaires (syllabes, phonèmes) et à l'inverse pour les méthodes phoniques. L'auteur a effectué une étude longitudinale auprès d'une population de 76 élèves suivis de la grande section de maternelle jusqu'au début de la 3ème année. Cette étude visait à évaluer l'impact de deux méthodes radicalement différentes sur la réussite en lecture estimée par la lecture de mots isolés et la compréhension de phrases écrites. L'une, nommée idéovisuelle, fait appel aux procédures orthographiques visant à la reconnaissance directe de mots écrits, à la mémorisation des formes orthographiques et à leur stockage dans le lexique mental. A l'inverse, la méthode phonique insiste sur l'établissement des CGP facilitant la reconnaissance de mots et vise à développer de manière spontanée et complémentaire la procédure orthographique (par entraînement à la mémorisation et à la discrimination visuelle). Les résultats indiquent un effet de la méthode. Au bout de deux ans, les performances des élèves en début de CE2 qui ont suivi la méthode phonique sont significativement supérieures en compréhension de texte et en reconnaissance de mots. Par contre, les performances des enfants bénéficiant d'un enseignement audio-visuel sont, non seulement systématiquement inférieures à celles de enfants bénéficiant de la méthode phonique, mais elles sont également significativement inférieures à la moyenne nationale. Ainsi, un enseignement explicite du code alphabétique est absolument indispensable pour développer une lecture efficace et experte.

Par ailleurs, plusieurs études montrent une influence de la méthode sur les capacités de segmentation de la parole qui se développent plus rapidement dans le cadre d'une méthode phonique. L'étude menée par Alegria, Pignot & Morais (1982) porte sur des enfants scolarisés en 1ère année du primaire qui apprennent à lire avec une méthode globale ou phonique. Différentes épreuves leur sont proposées : des épreuves de conscience phonémique et une épreuve de lecture de mots de un à cinq syllabes. Les résultats montrent que, quelle que soit la période, ce sont les enfants du groupe « méthode phonique » qui obtiennent les meilleurs scores à l'épreuve de lecture et aux épreuves de conscience phonémique. Les auteurs observent que des enfants de 1ère année d'une classe globale avaient des performances

beaucoup plus faibles que celles d'une classe phonique dans une épreuve consistant à inverser les phonèmes dans des mots monosyllabiques (os devient seau) présentés oralement, alors que les deux groupes étaient comparables lorsque la tâche consistait à inverser des syllabes (radis devient dira).

Certains auteurs (par exemple Vellutino & Scanlon, 2002) pensent que les méthodes globales et analytiques devraient être utilisées de façon complémentaire si l'on désire favoriser un développement optimal de la lecture. Le lecteur expert devrait être capable de passer d'une technique à l'autre selon le type de texte qu'il rencontre. Goswami (1986) propose un enseignement de la lecture basé sur l'utilisation d'analogies. Ainsi, deux méthodes d'enseignement de type phonique s'affrontent, une méthode phonique qui donne la priorité aux correspondances graphophonologiques (CGP) et une nouvelle méthode qui insiste sur l'analogie. Pour Goswami, les deux façons d'enseigner la lecture doivent se combiner. Il s'agirait d'entraîner les enfants à réaliser des analogies sur des unités larges afin de leur faciliter la découverte des unités réduites (les phonèmes) servant de base à l'apprentissage des CGP.

Dans l'ensemble, les résultats montrent un effet des méthodes d'enseignement au début du cursus scolaire. En revanche, l'information manque totalement en ce qui concerne la suite du développement. Comme le soulignent Content & Leybeart (1992) « *il ne semble pas que la méthode d'apprentissage de la lecture détermine des différences importantes dans les stratégies de lecture utilisées chez le lecteur habile, ni dans le chemin que l'enfant parcourt pour parvenir à la maîtrise de cette performance* » (p. 209).

S'il est important d'apprécier l'effet des méthodes d'enseignement de la lecture, il n'est pas moins intéressant de s'interroger sur les effets des caractéristiques linguistiques de la langue d'apprentissage.

7. Effets des caractéristiques linguistiques

Il existe dans le monde différents types d'écriture qu'on peut répartir en deux grandes classes : les écritures de type logographique ou les symboles graphiques se trouvent reliés aux morphèmes de la langue (l'écriture idéographique chinoise par exemple) et les écritures phonographiques où les symboles graphiques sont reliés selon les cas, soit aux phonèmes dans les écritures alphabétiques, soit aux syllabes de la langue dans les écritures syllabiques telles

que le *kana* japonais (Ecalle & Magnan, 2002 ; Peerman, 1992). Le point commun aux écritures alphabétiques et syllabiques est que toutes les deux représentent plus ou moins directement certaines caractéristiques sub-morphémiques du langage parlé (Peerman, 1992). Les écritures alphabétiques diffèrent toutefois entre elles par la régularité des correspondances grapho-phonologiques (CGP), ce qui signifie qu'il existe une relation entre les caractéristiques phonologiques de la langue parlée et la représentation graphique. Par exemple, dans la langue française, le son /a/ se représente généralement par la lettre [a], et le groupe de lettres [*oi*] se prononce généralement /wa/ comme dans *oiseau*. Il est ainsi possible de décrire un ensemble de règles pouvant prédire la prononciation de lettres ou de groupe de lettres.

Selon Alegria, Marin, Carrillo & Mousty (2003), les systèmes orthographiques peuvent être classés en « superficiel » et « profond » suivant le niveau de langue qu'ils représentent. L'espagnol et l'allemand, par exemple, possèdent des systèmes qualifiés de superficiels parce qu'ils reflètent de manière relativement fidèle la phonologie de surface. Le français et l'anglais, par contre, font partie des systèmes qualifiés de profonds parce qu'ils se proposent de représenter la langue au niveau phonologique, morphologique et syntaxique. Ainsi, en français, qui représente systématiquement la morphologie, certains graphèmes ne représentent aucun phonèmes (par exemple, la phrase *ils marchent* comporte des marques du pluriel qui n'apparaissent pas dans la prononciation). Lorsqu'on envisage la relation entre la représentation écrite des mots et leur prononciation, il est habituel de considérer l'orthographe française plus régulière que l'orthographe anglaise. Ceci signifie que les lecteurs français recouraient plus aux processus d'assemblage phonologique que les lecteurs anglais. De même, la grande régularité des CGP en espagnol devrait encourager le lecteur à exploiter davantage les processus d'assemblage phonologique que les lecteurs anglais et français. Plusieurs études ont abordé ce problème en comparant les mécanismes d'acquisition de l'orthographe dans des systèmes présentant des différences importantes au niveau du rôle potentiel du code phonologique. Les études comparatives entre écritures permettent, d'une part, de constater si la manière dont l'écriture encode le langage parlé détermine une différence quant à la fréquence d'exploitation des processus d'assemblage et d'autre part, d'examiner le degré d'universalité des résultats obtenus avec l'écriture anglaise.

Des études effectuées dans diverses langues montrent l'universalité de certains processus de base en particulier la maîtrise de l'assemblage phonologique quelle que soit la langue objet d'apprentissage (Boukadida, 2008 ; Ecalle & Magnan, 2007a, Ferrand, 1995 ; Gombert,

1992b ; Wimmer & al, 1991). Cependant, les unités phonologiques utilisées selon les langues sont variables et la nature des difficultés des apprentis lecteurs semble varier d'une langue à l'autre (Cossu, Shankweiler, Liberman & Gugliotta, 1995 ; Shankweiler, Crain, Brady & Maracuso, 1992). La mise en œuvre de la procédure sublexicale ainsi que son fonctionnement devraient dépendre des spécificités linguistiques de chaque langue. L'apprentissage de la lecture dépend de la transparence des correspondances grapho-phonémiques, qui est très élevée dans certaines langues (espagnol) et très faibles dans d'autres (anglais). Ainsi, la procédure lexiacale est davantage utilisée quand l'orthographe est opaque, probablement pour compenser les difficultés de mise en œuvre du décodage. Une grande partie des recherches sur l'acquisition de la lecture porte sur l'observation d'enfants anglophones. L'anglais présente un système orthographique particulier dû à l'inconsistance des relations graphèmes-phonèmes. Ainsi, la faible exploitation des processus d'assemblage phonologique chez les lecteurs anglophones pourraient résulter de l'existence d'un grand nombre d'irrégularités graphophonologiques dans la langue (Peerman, 1991).

Les différences entre langues, concernant les déterminants phonologiques engagés dans la reconnaissance des mots écrits, entrainent des difficultés différentes chez les apprentis lecteurs. Ainsi, l'étude de Duncan & al. (2006, cité par Ecalle & Magnan, 2007a) montre que la présence de frontières syllabiques claires en français contribue à l'émergence d'une conscience syllabique plus précoce. L'étude a porté sur des pré- et apprentis lecteurs français et anglais. Les résultats indiquent que les enfants anglais, avant l'enseignement formel de la lecture, ont des performances dans les tâches de manipulation de syllabes bien inférieures à celles de leurs homologues français. D'autres études (Bruck, Genesee & Caravolas, 1997 ; Goswami, Gombert & Barrera, 1998 ; Frith, Mann & Landerl, 1998 ; Mann & Wimmer, 2002) ont été réalisées afin d'examiner si, dans la reconnaissance des mots, la régularité des relations entre orthographe et unités phonologiques favoriserait l'utilisation d'un code phonologique assemblé[30]. Goswami, Gombert & Barrera (1998) ont conduit une étude interlangue avec des enfants français, anglais et espagnols de 7, 8 et 9 ans. Ces enfants devraient lire des pseudo-mots analogues partageant la même rime orthographique et phonologique avec les mots dont ils étaient issus (par exemple *dake* et *bicket* sont issus de *cake* et *ticket* en anglais). La lecture de ces items a été comparée à celle de pseudo-mots non analogues (par exemple *faish* et *terket* en anglais). Pour chaque tranche d'âge, les enfants ont

[30] Nous allons décrire les études comparatives interlangues qui ont essayé de détermeiner les meilleurs prédicteurs de la lecture dans différentes langues (anglais, français, allemand, espagnol, arabe) dans le chapitre 4.

été appariés sur leur niveau de lecture et sur leur connaissance des mots utilisés pour construire les pseudo-mots analogues. Dans l'ensemble, les scores les plus élevés s'observent chez les enfants espagnols (89%). Ceux des français sont supérieurs à ceux des anglais (72% *vs* 52%). Les résultats indiquent également de fortes différences entre les trois groupes linguistiques pour les effets d'analogie et de longueur syllabique. Le bénéfice de l'analogie est très fort chez les anglophones et les francophones alors que l'effet de la longueur syllabique est observé chez les hispanophones. La comparaison des résultats en lecture de mots et de pseudo-mots révèle une faiblesse des scores des anglophones comparativement aux autres groupes, surtout en lecture de pseudo-mots. Les auteurs montrent également que la présence de voisins orthographiques réduit le temps et les erreurs de lecture de pseudo-mots chez des enfants de 1ère année primaire en français et en anglais. En espagnol, par contre, les effets des voisins orthographiques s'observent uniquement sur le temps de lecture. La simplicité du code alphabétique espagnol semble donc réduire l'influence des voisins orthographiques de la lecture.

Dans une optique similaire, Frith, Wimmer & Landerl (1998) ont mené une étude avec des enfants anglophones et germanophones de même âge qui devraient lire des mots appariés en difficulté orthographique et en familiarité («summer» / «sommer») et des pseudo-mots dérivés des mêmes mots («rummer» / «rommer»). Dans une seconde expérience, la fréquence et la lexicalité ont été manipulées. Les performances des anglophones sont inférieures à celles des germanophones en lecture de mots (80% *vs* 95%) et surtout en lecture de pseudo-mots (59% *vs* 88%). Pour les mots fréquents, les performances des enfants de 8 ans des deux groupes linguistiques sont similaires alors que les anglophones manifestent de moins bons scores pour les mots rares. Les résultats ont également permis de constater que les erreurs des anglophones concernaient majoritairement les voyelles, mais pas les germanophones. Ceci peut s'expliquer par la plus grande consistance de l'orthographe de l'allemand. L'ensemble de ces résultats indique que les deux groupes linguistiques n'utilisent pas exactement les mêmes procédures de lecture. L'absence d'effet facilitateur de la fréquence et de l'homophonie en lecture à haute voix indique que les germanophones s'appuieraient essentiellement sur la procédure grapho-phonologique pour lire. Les enfants anglophones utilisent davantage la procédure lexicale pour compenser les difficultés de mise en œuvre de la procédure sublexicale.

Les facteurs linguistiques induisent aussi des différences dans les unités de lecture utilisées. La moindre consistance des unités grapho-phonémiques en anglais semble favoriser un plus

grand recours à des unités plus large que le phonème, en particulier la rime. En revanche, la prise en compte de la rime n'a pas d'incidence majeure dans les langues qui comportent des voyelles clairement articulées, comme en français ou en espagnol.

8. Le poids des facteurs sociaux

L'apprentissage de la langue écrite constitue un véritable enjeu de société et fait l'objet d'un grand nombre de recherches relevant de disciplines diverses (psychologie cognitive, psycholinguistique, didactique, socio-psychologie), chacune se centrant sur une dimension particulière. Les études socio-psychologiques ont mis en évidence le rôle déterminant du milieu familial, plus particulièrement, des représentations, des attitudes ou encore des pratiques familiales vis-à-vis des apprentissages de l'école (Ben Fadhel, 2003 ; Carayon, 1991).

8.1. La représentation sociale sur la lecture

La lecture, à la fois outil de communication et objet social, met en rapport les individus et s'inscrit dans des contextes qui génèrent des attitudes, des positions et des pratiques, définissant ainsi un *habitus lectural* (Ecalle & Magnan, 2002). C'est dans ces interactions complexes classe sociale/milieu familial/intelligence que se constitue l'habitus lectural de l'enfant. La lecture et son acquisition font très tôt l'objet d'un ensemble de représentation en cours de développement chez le jeune enfant, en particulier au moment de l'apprentissage proprement dit de la lecture. Pour étudier les liens entre la compétence lexique et les représentations sociales de la lecture, Carayon (1991) propose aux enfants des images représentant trente situations de lecture différentes (information, jeu, loisir, utilitaire). Les enfants doivent dire si la situation présentée est une situation de lecture et justifier leurs réponses. L'auteur distingue trois types de représentations : une représentation défavorable à la lecture, une représentation favorable et une représentation ambivalente. Les sujets défavorables à la lecture envisagent celle-ci comme une activité imposée qui ne fait partie de leurs pratiques. La représentation favorable est liée à l'idée de plaisir, de loisir où la lecture est une activité courante intégrée à la vie quotidienne. Dans la représentation ambivalente, la lecture est considérée comme une source d'information contribuant, dans le futur, à la réussite sociale. Les résultats montrent que les enfants ayant une représentation favorable ont de bonnes performances en lecture, ceux qui portent une représentation défavorable ont un faible niveau de lecture. Les sujets à la position ambivalente ne se caractérisent par aucune

performance spécifique[31]. Cela confirme la nature psycho-sociale de l'apprentissage de l'écrit, qui ne peut être réduit à une simple technique scolaire.

8.2. Les pratiques et les conceptions éducatives des parents

Les parcours scolaires dépendent de nombreux facteurs, certains d'entre eux étant directement associés à l'élève, avec principalement ses capacités à apprendre et à s'adapter au monde scolaire. D'autres facteurs concernent l'environnement proche de l'enfant, ses conditions de vie au sein de sa famille ou encore l'aide apportée par les parents au niveau de la scolarité. L'enfant évolue dans un dispositif socio-institutionnel, l'école, où des attentes et des représentations réciproques déterminent la relation enseignant-enseigné (Ben Fadhel, 2003). Les conditions de vie et de travail liées au statut socio-économique des parents déterminent leurs pratiques éducatives qui, à leur tour, agissent sur le développement intellectuel de l'enfant. A ce titre, les caractéristiques du contexte éducatif et les interactions entre l'enfant et ce contexte s'avèrent des composantes à prendre en compte pour une meilleure connaissance du fonctionnement cognitif de l'apprenti lecteur. L'effort cognitif sera plus facilement réalisé si l'enfant se trouve dans un contexte éducatif où l'écrit occupe une place privilégiée. L'école elle-même, par son organisation les conditions et la qualité de l'enseignement reçu, exerce aussi une influence notable sur les parcours des élèves.

L'élève construit son rapport au savoir scolaire non seulement au sein du cercle restreint de l'école mais aussi à travers la médiation familiale. Cette médiation assure la socialisation linguistique de l'enfant et participe activement à l'apprentissage formel de l'écrit au sein même de l'école (Ben Fadhel, 2003). Ben Fadhel (2003) analyse les conceptions et les pratiques éducatives des parents de 74 enfants du cours préparatoire à partir d'un questionnaire composé de trois rubriques principales : les pratiques déclarées des parents autour des activités de lecture-écriture chez l'enfant, les représentations des parents sur la question de l'enseignement-apprentissage de la lecture-écriture et les valeurs éducatives que les parents attendent de l'école. Les résultats indiquent que le rapport à l'écrit des parents, nettement marqué par le milieu social et culturel, est un facteur déterminant des compétences des enfants dans ce domaine. Ce sont ainsi les enfants les plus initiés dans leur milieu familial à des activités fonctionnelles de lecture-écriture qui développent les stratégies d'apprentissage les plus efficaces et qui expriment le rapport le plus autonome aux savoirs de l'école. La

[31] D'autres travaux montrent que la structure de la représentation sociale d'enfants non-lecteurs apparait comme inorganisée, floue au contraire de celle bien organisée des faibles et des bons lecteurs (Courtial & Ecalle, 1999).

lecture partagée d'histoires entre adultes et enfants constitue une routine interactive idéale à l'acquisition du langage oral et peut amener l'enfant à engager un travail de réflexion sur l'écrit. L'activité de lecture sera intégrée dans le cercle des pratiques quotidiennes de l'enfant et aura un statut autre que celui d'une discipline scolaire obligatoire. D'autres expériences familiales partagées comme les comptines et les chansons participent indirectement aux connaissances liées à l'écrit par le développement des habiletés phonologiques, indispensables à l'apprentissage de la lecture. Par conséquent, les pratiques pédagogiques devraient s'intéresser davantage à cette dimension relationnelle en favorisant l'articulation des exercices scolaires avec les expériences familiales.

9. Conclusion

La lecture est une activité complexe qui fait appel à de nombreuses compétences, qu'elles soient linguistiques ou non. De fait, elle implique la mise en œuvre d'opérations de différentes natures mettant en jeu des habiletés spécifiques au traitement de l'information écrite et des habiletés générales. Avant même le début de l'école primaire, l'enfant dispose de connaissances conscientes et d'habiletés implicites qui constitueront le socle sur lequel va s'appuyer l'enseignement de la lecure et de l'écriture (Gombert & al., 2004). Parmi les découvertes les plus saillantes de ces trente dernières années, il faut signaler le consensus qui s'est formé autour de l'importance des habiletés phonologiques dans l'acquisition de la lecture (Demont & Gombert, 2007 ; Ecalle & Magnan ; 2007a ; 2007b ; Gombert, 1990 ; Goswami & Bryant, 1990 ; Sprenger-Charolles & Casalis, 1996). En effet, un des problèmes principaux pour les enfants pré-lecteurs est de comprendre que la parole peut être segmentée en unités de plus en plus petites jusqu'aux phonèmes, et que ces unités élémentaires sont celles qui sont représentées par les lettres écrites. La capacité des enfants d'âge préscolaire à compter les syllabes d'un mot ou à détecter des rimes a été souvent mise en évidence (Gombert & Colé, 2000). Il s'agit là de l'intervention de processus « épilinguistiques » au sens où les définit Gombert (1990) c'est-à-dire de comportements qui ne relèvent pas d'un contrôle conscient qui serait exercé par l'enfant à partir de ses connaissances sur ces propres opérations de traitement linguistique (ce qui relèverait d'une activité métalinguistique).

L'élève doit commencer par apprendre le principe général du code alphabétique, puis acquérir un nombre suffisant de correspondances entre les graphèmes et les phonèmes pour commencer à décoder. Cette correspondance graphème-phonème va rapidement et progressivement devenir automatique, jouant un rôle moteur et génératif dans l'apprentissage

de la lecture. Ainsi, certaines habiletés phonologiques se développent précocement, de façon implicite, en lien avec l'expérience langagière, tandis que d'autres apparaissent plus tardivement, sous l'effet de l'instruction formelle. Casalis & Lecoq (1992) considèrent que, pour aborder la lecture, l'enfant doit être doté d'un système de représentation et de traitement de l'information phonologique (SRT). Avant l'apprentissage de la lecture, ce système ne lui permet de traiter que les unités phonologiques les plus larges (rimes, syllabes). Au contact avec l'écrit, ce système se développe et peut traiter des unités plus abstraites, comme les phonèmes.

Les études montrent, dans l'ensemble, que les capacités métaphonologiques entretiennent des relations fortes et spécifiques avec l'acquisition de la lecture dans un système alphabétique. En particulier, ces capacités permettent de distinguer des populations de bons et de faibles lecteurs, elles sont de bons prédicteurs du niveau ultérieur en lecture et sont améliorables par des exercices spécifiques, les entraînements phonologiques, facilitant l'acquisition de la lecture. Un autre aspect important de la conscience phonologique est son rapport avec la connaissance des lettres. En effet, les enfants commencent à relier l'oral et l'écrit sur la base du nom des lettres. La connaissance du nom des lettres stimulerait le développement de la conscience phonémique des enfants au tout début de l'apprentissage de la lecture (Ecalle & al., 2008 ; 2009 ; Biot-chevrier & al., 2008 ; Foulin, 2005).

Toutefois, on ne peut pas réduire l'apprentissage de la lecture à la maîtrise de la structure segmentale du langage écrit. Certains facteurs sont largement négligés, c'est le cas par exemple des connaissances morphologiques et des processus cognitifs. Plusieurs recherches ont montré, sur le modèle de ce que l'on connait des liens entre la dimension phonologique et l'apprentissage de la lecture, l'importance de la dimension morphologique (Boukadida, 2008 ; Gombert, 2003 ; Marec-Breton, 2009 ; 2003). La structure morphologique constituerait une aide notable en lecture en appui du traitement graphophonologique. Ainsi, dès les toutes premières étapes de l'apprentissage de la lecture, les enfants disposent de connaissances morphologiques implicites et explicites non négligeables qui jouent un rôle important dans les procédures d'identification de mots écrits (Boukadida & al., 2009 ; Marec-Breton & al., 2005). Différentes études (Carlisle, 1995 ; Lecoq & al., 1996) indiquent que les connaissances morphologiques affectent la compréhension. D'autres montrent que les ces connaissances affectent les performances aux tâches de lecture de mots (Carlisle & Nomanbhoy, 1993). Enfin, certaines études montrent que les connaissances morphologiques affectent les performances en lecture de mots et en compréhension (Casalis & Louis-Alexandre, 2000).

La notion de trouble de l'empan visuo-attentionnel, récemment formulée, suggère quant à elle qu'une difficulté à traiter en parallèle les lettres de la séquence du mot pourrait être à l'origine de certaines formes de troubles dyslexiques, indépendamment des capacités de traitement phonologique des enfants. Par ailleurs, certaines études antérieures montrent que les facteurs langagiers et cognitifs chez le jeune enfant ne suffisent pas pour expliquer les différences individuelles dans l'acquisition des compétences relatives à la langue écrite. Les études interlangues ont permis de rendre compte de l'incidence de certaines caractéristiques de la langue dans laquelle s'effectue l'apprentissage sur les procédures de lecture. En particulier, la régularité des correspondances grapho-phonologiques qui se traduit par une grande transparence de l'orthographe induit une utilisation préférentielle de la médiation phonologique et facilite le décodage. Ainsi, lorsque l'orthographe est transparente, les informations phonologiques seraient activées systématiquement. En revanche, lorsque l'orthographe est opaque, comme en anglais, les apprentis lecteurs s'appuient davantage sur les informations lexicales. En ce qui concerne la langue arabe, la structure morpho-syntaxique des mots d'une part et le double système d'écriture d'autre part pourraient avoir un impact sur les stratégies d'identification des mots. La langue arabe se différencie de l'anglais à plusieurs niveaux qui ont des répercussions sur les procédures de traitement utilisés par les lecteurs de ces deux langues.

Les connaissances issues d'approches complémentaires viennent enrichir et compléter les modèles théoriques de la psychologie cognitive en soulignant notamment l'influence non négligeable des variables conatives (Guimard & Florin, 2008). Il est évident, en effet, que tous les enfants ne sont pas intéressés au même degré par l'apprentissage et l'usage de la lecture, qu'ils ne tirent pas tous de cette activité les mêmes satisfactions. Ces différences conatives peuvent intervenir non seulement pour faire varier d'un enfant à un autre le temps et l'effort investis dans la lecture, mais aussi pour faire varier les processus mêmes, les stratégies, qu'un lecteur met en œuvre (Reuchlin, 1990). D'autres travaux ont examiné le rôle du contexte éducatif dans les différents processus impliqués dans l'activité de lecture. Ils suggèrent que les enfants fortement exposés à l'écrit élaborent implicitement une sensibilité aux unités orthographiques composant les mots de la langue écrite (Ecalle & Magnan, 2002).

Chapitre 4:Prédicteurs de l'apprentissage de la lecture

Avant l'apprentissage explicite de la lecture, l'enfant construit un ensemble de connaissances implicites et explicites sur la langue écrite (connaissances phonologiques et morphologiques, connaissances des lettres et connaissances sur les régularités orthographiques) qui constituent des précurseurs dont on a pu montrer que leur absence entravait sérieusement l'apprentissage ultérieur de la lecture. Les connaissances enseignées seront, en effet, assimilées par chaque enfant en fonction des précurseurs qu'il a construits (Ecalle & Magnan, 2002). De ce point de vue, au moyen d'études longitudinales prédictives, qui examinent une population d'enfants avant l'apprentissage formel de la lecture, les chercheurs essaient de spécifier le pattern de déficience des futurs faibles lecteurs et de s'interroger sur la causalité de l'échec en lecture. Cette approche a permis d'identifier quelques prédicteurs de la réussite ou de l'échec futurs dans l'apprentissage de la lecture. L'objectif principal de ce chapitre est de présenter un ensemble de travaux ayant pour objet de déterminer les prédicteurs de réussite ou d'échec de l'apprentissage de la langue écrite.

1. Quelles habiletés précoces prédisent les compétences ultérieures de lecture ?

La période qui précède l'apprentissage explicite de la lecture est marquée par l'acquisition de certaines habiletés qui déterminent la façon dont les enfants perçoivent et traitent cet apprentissage. Les déficits relevés dans ce domaine avant l'apprentissage de la lecture doivent ainsi prédire la réussite ou l'échec à cet apprentissage. Différentes études permettent de valider cette hypothèse :

- les études longitudinales prédictives qui permettent de cerner le pouvoir prédictif de différentes compétences évaluées à un moment donné sur les compétences développées ultérieurement.
- les études comparatives interlangues qui permettent d'étudier le rôle de certaines contraintes linguistiques sur le processus de l'apprentissage de la lecture.

- les études comparatives qui rendent compte des performances de bons et de faibles lecteurs ou enfants dyslexiques. Ces élèves ont soit le même âge chronologique, soit un âge différent mais avec un niveau de lecture comparable.
- les études associées à un programme d'entraînement métalinguistique permettant de se prononcer sur le caractère causal des connaissances phonologiques et morphologqiues.

1.1. Etudes longitudinales prédictives

Les études prédictives apportent une contribution à la compréhension de la réussite ou de l'échec spécifiques en lecture[32].

Ellis et Large (1988) ont effectué une étude longitudinale portant sur 40 enfants suivi depuis le début de l'apprentissage, à l'âge de 5 ans, jusqu'à l'âge de 7 ans. Ces enfants sont testés chaque année pour mesurer différentes habiletés : compétences phonologiques, vocabulaire, mémoire à court terme, connaissances morphologiques, perception et discrimination visuelle et niveau intellectuel. L'objectif est de déterminer les habiletés précoces qui sont associées à l'apprentissage de la lecture. Les différentes analyses de corrélation révèlent que les prédicteurs les plus puissants des compétences ultérieures en lecture sont la conscience phonologique (segmentation syllabique et phonémique, inversion phonologique), la connaissance de l'alphabet (connaissance de lettres) et la mémoire à court terme. Cette étude montre que l'apprentissage des correspondances graphophonologiques ne prédit pas le niveau de lecture à 5 ou à 6 ans (r=.27, r=.13) mais à l'âge de 7 ans (r=.73). De même, les connaissances morphologiques et syntaxiques ne semblent pas prédire la réussite en lecture à l'âge de 5 ans mais sont de bons prédicteurs de l'habileté de lecture à 7 ans (respectivement r=.54 et r=.58). Ces données suggèrent que le début de lecture ne nécessite qu'un usage limité des connaissances morphologiques et syntaxiques, mais avec l'avancement dans la scolarité ces deux compétences se développent mutuellement. Selon les auteurs, les prédicteurs varient en fonction des différents stades de lecture. Ainsi, au cours du stade logographique, ce sont la conscience phonologique, la connaissance des lettres et la discrimination visuelle qui prédisent le niveau de lecture l'année suivante. Au stade alphabétique, les prédicteurs de de la lecture sont l'apprentissage des CGP et la mémoire à court terme. Enfin, le stade

[32] Les techniques statistiques utilisées sont des analyses de régression, dérivées des analyses de corrélation, où la variable à expliquer correspond aux performances dans un domaine de la langue écrite et les variables explicatives candidates aux performances à différentes épreuves verbales (métalinguistiques la plupart du temps) et des épreuves mesurant divers domaines.

orthographique implique la participation de la conscience phonologique et des connaissances morphologiques et syntaxiques.

Muter & al. (1998) ont suivi 38 enfants anglophones depuis la maternelle jusqu'à la 2ème année primaire. Au temps 1, les enfants ont passé une version réduite du WISC et une batterie de tests phonologiques comportant une partie « rime » (tâche de détection et de production de rimes) et une partie « segmentation » (tâche d'identification et de suppression de phonèmes) ainsi qu'une épreuve de connaissance des lettres où l'enfant devait nommer les 26 lettres de l'alphabet présentées dans un ordre aléatoire. Aux temps 2 et 3, deux épreuves de lecture de mots (mots isolés et mots dans des phrases) et une épreuve de production orthographique de mots isolés sont proposées. Les résultats montrent que les scores en lecture au temps 2 sont expliqués (63% de la variance totale), par ordre d'importance décroissante, par les connaissances des lettres, les scores en segmentation phonologique et par les scores composites associant « connaissance des lettres et segmentation ». Au temps 3, les scores en lecture de mots sont expliqués par ceux réalisés un an plus tôt en lecture et par les connaissances des lettres au temps 3.

Kirby, Parrila & Pfeiffer (2003) ont essayé de déterminer le rôle de la conscience phonologique et de la vitesse de dénomination dans le développement de la lecture. L'objectif était de savoir comment ces deux compétences, mesurées avant l'apprentissage formel de la lecture, peuvent prédire le futur niveau de lecture. L'étude a porté sur 161 enfants suivis longitudinalement depuis la grande section de maternelle (âge moyen 5 ans 5 mois) jusqu'à la 5ème année primaire. En maternelle, 4 épreuves de conscience phonologique ont été administrées aux enfants : une tâche de segmentation phonémique, une tâche de suppression phonémique, une tâche de fusion phonémique et une tâche de fusion attaque-rime (on présente à l'enfant l'attaque et la rime d'un mot et l'enfant doit prononcer le mot qui en résulte). Deux épreuves de dénomination sont également administrées aux enfants : dénomination de couleurs et dénomination d'images. Dans chaque tâche, l'enfant est amené à nommer une série de 32 images ou couleurs. Les auteurs ont effectué également des mesures des habiletés mentales générales et une mesure de connaissance des lettres. Le niveau de lecture a été évalué systématiquement chaque année. Les analyses de régression mettent en évidence que le niveau de conscience phonologique et la vitesse de dénomination, mesurés en maternelle, contribuent indépendamment à la prédiction du niveau ultérieur en lecture. La conscience phonologique constitue le prédicteur le plus puissant en maternelle et en 1ère année. En revanche, la vitesse de dénomination a un effet minime au début de l'apprentissage

mais sa valeur prédictive augmente avec l'avancement dans la scolarité (5ème année). Alors que la conscience phonologique est associée aux performances en identification de mots, la vitesse de dénomination semble plus reliée à la lecture-compréhension. En effet, la compréhension dépend de l'intégration des informations présentes en mémoire de travail. Les enfants ayant une vitesse de dénomination lente auront des difficultés à comprendre les items les plus difficiles. Les résultats indiquent également que les enfants ayant le niveau le plus bas en conscience phonologique et en dénomination rapide en maternelle sont ceux qui progressent le plus lentement dans l'acquisition de la lecture et sont ceux qui souffrent le plus de difficultés de lecture en 5ème année primaire. Cet effet est observé aussi bien pour l'identification des mots que pour la lecture-compréhension.

L'étude de Schatschneider, Carlson, & Foorman (2004) avait pour objectif de déterminer la valeur prédictive de différentes mesures réalisées en maternelle (GSM) sur les performances en lecture au grade 1 et au grade 2. En GSM, différentes compétences jugées importantes dans le développement des performances en lecture ont été évaluées : conscience phonologique, connaissance des lettres, dénomination rapide de lettres et d'objets, vocabulaire et compétences visuo-motrices. A la fin des grades 1 et 2, la performance en lecture a été évaluée à travers différentes tâches : identification de mots et de lettres, vitesse de lecture de mots. Les résultats indiquent que trois variables expliquent une part de variance unique : la conscience phonologique, la connaissance du son des lettres et la dénomination rapide de lettres. Au début de GSM, la connaissance du nom des lettres et la dénomination rapide d'objets constituent également des prédicteurs puissants de la performance en lecture. En revanche, leur valeur prédictive diminue à la fin de la GSM et du grade 1. Les mesures de vocabulaire, de syntaxe et des compétences visuo-motrices n'expliquent pas une part significative de la variance en lecture. En comparant les résultats obtenus au grade 1 et au grade 2, les auteurs ont mis en évidence une diminution des relations entre les prédicteurs de GSM et les performances en lecture au grade 2 mais aucune différence n'est observée en ce qui concerne le pattern des variables contribuant à ces performances.

Muter, Hulme, Snowling & Stevenson (2004) ont mené une étude longitudinale avec 90 enfants britanniques âgés de 4 ans 2 mois à 5 ans 2 mois suivis pendant deux ans. L'objectif est d'explorer la relation entre certaines habiletés précoces (compétences phonologiques, connaissance des lettres, connaissances morphologiques et connaissance du vocabulaire) et le niveau ultérieur en identification de mots et en compréhension durant les deux premières années de l'acquisition de la lecture. Les enfants ont reçu une batterie de tests en trois temps

pendant une période de deux ans. Au temps 1, différentes épreuves phonologiques sont administrées aux enfants : détection et production de rimes, suppression phonologique, inversion phonologique, ainsi qu'une épreuve de connaissance des lettres et un test de lecture. Au temps 2, trois mesures sont ajoutées aux épreuves administrées en temps 1 : une épreuve syntaxique, une épreuve de morphologie et un test de lecture de mots. Enfin, les enfants ont passé différents tests de lecture mesurant les compétences en identification de mots et en compréhension. Les auteurs ont eu recours à la modélisation par les équations structurelles pour étudier les liens prédictifs entre les différentes variables. Les résultats montrent que le niveau d'identification de mots au temps 3 est prédit par trois variables qui expliquent à elles seules 88% de la variance : l'identification de mots au temps 2, la connaissance des lettres et la sensibilité phonémique. En revanche, la sensibilité à la rime, le niveau de vocabulaire et les connaissances morphologiques ne sont pas aussi importants pour l'identification de mots mais semblent prédire de façon significative la lecture-compréhension.

Les résultats des études impliquant des enfants francophones sont largement similaires à ceux relevés avec des enfants anglophones. Lecoq (1991) a réalisé une étude longitudinale auprès d'une large population d'enfants de la GSM suivis jusqu'en CE1. L'objectif principal était de procéder à un examen aussi précoce que possible des domaines de compétences qui étaient, selon la littérature, les plus prédictifs de la réussite ou de l'échec de l'apprentissage de la lecture. L'étude a porté sur une multitude d'épreuves explorant la conscience phonologique, divers aspects de la mémoire et la vitesse de dénomination rapide au cours du développement et leurs relations avec les performances ultérieures en lecture. Les tâches phonologiques sollicitent des opérations d'identification, de comptage, de catégorisation, de segmentation, de soustraction ou de fusion. Les épreuves de mémoire visaient à recueillir des informations sur l'empan mnésique en fonction de certaines caractéristiques des items, en l'occurrence, leur longueur, leur familiarité, leur similarité phonémique ou sémantique et leur catégorie (chiffres ou mots). Dans les épreuves de dénomination rapide, une série d'images que le sujet doit dénommer est présentée. L'objectif consiste à vérifier la précision et la rapidité d'accès au lexique. Les capacités de lecture des enfants ont été évaluées par deux tests de lecture. Les résultats indiquent que l'épreuve la plus prédictive, quel que soit le test de lecture utilisé, implique la discrimination de phonèmes (trouver le mot qui ne commence pas par la même consonne que la cible). Toutefois, la valeur prédictive de la mémoire et de la rapidité de dénomination est beaucoup moins puissante que celle des compétences phonémiques

précoces. Le QI non verbal ne ressort pas comme variable explicative du niveau ultérieur de lecture.

L'étude longitudinale de Demont & Gombert (1996) a consisté en un suivi d'enfants français de la dernière année d'école maternelle (GSM) à la fin de la 3ème année d'école élémentaire (CE2) afin de déterminer la contribution respective de la conscience phonologique et de la conscience syntaxique à l'efficience ultérieure en lecture. Différentes épreuves de conscience linguistique exigeant une activité réflexive sur le langage, soit sur son aspect phonologique, soit sur son aspect syntaxique, ont été présentées aux enfants lors de passations individuelles. L'efficience en reconnaissance de mots et en compréhension était simultanément évaluée. Les analyses de régressions multiples ont permis de mettre en évidence la contribution et la valeur prédictive des compétences métalinguistiques aux capacités ultérieures en lecture. La conscience phonologique et la conscience syntaxique apparaissent contribuer conjointement à l'apprentissage de la lecture. Leur contribution apparaît toutefois moins massive à l'explication de la variance des capacités de compréhension. Par contre, les capacités à gérer intentionnellement l'ordre des mots (évaluées à l'aide des épreuves impliquant l'inversion de l'ordre des constituants) contribuent uniquement aux capacités en compréhension. Il existerait ainsi une relation spécifique entre conscience syntaxique et compréhension.

L'étude de Piquard-kipffer (2003) a porté sur 85 enfants suivis de la grande section de maternelle (GSM) à la fin de la 2ème année du primaire (de 4 à 8 ans). Les enfants ont passé différentes épreuves : épreuves phonologiques, épreuve de connaissance du nom et du son des lettres, test de vocabulaire, épreuve de mémoire verbale, une épreuve de vitesse d'articulation et une épreuve de compréhension et un test d'intelligence non verbale. L'analyse de régressions indique que les variables qui possèdent le plus fort pouvoir prédictif sont, par ordre décroissant, la rapidité de dénomination, les capacités de discrimination phonémique, celles de segmentation phonémique, puis la mémoire verbale. Ni les capacités de mémoire visuelle, ni celles de compréhension orale et de vitesse d'articulation ne prédisent le niveau ultérieur en lecture. Deux d'entre elles (les capacités de discrimination phonémique et de dénomination rapide) expliquent 43,5% de la variance en lecture en fin de 2ème année et permettent à elles seules de classer 83,5% des enfants en fonction de leur niveau de lecture ultérieur.

Abu-Rabia (2007) a effectué une étude longitudinale afin de tester la valeur prédictive des compétences phonologiques et morphologiques sur l'identification de mots isolés et la

compréhension en langue arabe chez des enfants normolecteurs et des enfants dyslexiques d'âges différents (grades 3, 6, 9 et 12). Les analyses de régression révèlent que la morphologie est le prédicteur le plus puissant des performances ultérieures en identification de mots isolés et en lecture-compréhension aussi bien chez les normolecteurs que chez les lecteurs dyslexiques à tous les niveaux scolaires. Les connaissances morphologiques (identification et production) est un déterminent crucial de l'apprentissage de la lecture quel que soit l'âge ou le niveau de lecture (Abu-Rabia & al. 2003).

L'étude longitudinale réalisée par Boukadia & al. (2009) confirme ces résultats. L'étude a porté sur une quarantaine d'enfants tunisiens suivis de la fin de 1ère année à la fin de 2ème anné parimaire. Pour l'évaluation des connaissances phonologiques et morphologiques, les auteurs ont utilisé quatre tâches : détection d'intrus de rime, suppression phonémique, reconnaissance de l'appartenance à la même famille morphologique et segmentation morphémique. Les analyses corrélationnelles et de régressions multiples ont montré que les connaissances phonologiques et morphologiques interviennent dans l'explication de la variance en lecture. En 1ère année, les résultats indiquent que toutes les tâches sont corrélées au niveau de lecture, excepté la tâche de détection d'intrus de rime. En 2ème année, la tâche de reconnaissance de l'appartenance à la même famille morphologique présente le coefficient de corrélation le plus élevé avec le niveau de lecture (r =.62). Les performances à la tâche de suppression phonémique, en 1ère année, expliqueraient 40 % de la variance en lecture une année plus tard. Les connaissances morphologiques interviendraient également dans l'explication des scores de lecture (10% de la variance en lecture). La part de contribution de la dimension morphologique augmente entre la 1ère année et la 2ème année.

1.2. Etudes comparatives interlangues

Les études interlangues visent une meilleure compréhension de l'articulation entre le fonctionnement d'une langue et les stratégies mises en œuvre par l'apprenant pour l'acquérir afin d'étudier le rôle de certaines contraintes linguistiques sur le processus de l'apprentissage de la lecture (Morais, 1995 ; Peerman, 1991). Les premiers travaux interlangues qui ont porté sur des langues alphabétiques latines ont montré que la facilité et la vitesse d'apprentissage était liée à la régularité et à la transparence des relations entre graphèmes et phonèmes. Dans certaines langues (l'espagnol par exemple), ces relations sont plutôt transparentes. Il est donc possible d'obtenir la prononciation des mots à partir de la connaissance des graphèmes. Dans d'autres langues, l'existence d'un grand nombre d'irrégularité orthographiques engendre une

faible exploitation des processus d'assemblage phonologique. Ceci pourrait amener l'apprenti lecteur à abandonner progressivement le recours à cette procédure pour identifier les mots.

Dans une étude interlangue, Bruck, Genesee & Caravolas (1997) ont suivi des enfants canadiens francophones et anglophones. Les habiletés métaphonologiques ont été évaluées en milieu de grande section de maternelle (GSM) par six épreuves de jugement de similarité sur des paires de pseudo-mots. Une épreuve de comptage syllabique pour des pseudo-mots a également été administrée ainsi qu'un test de suppression du premier phonème dans des pseudo-mots et une épreuve de connaissance du nom des lettres de l'alphabet. Les performances en lecture de mots et de pseudo-mots ont été examinées à la fin de la 1ère année. En ce qui concerne les épreuves métaphonologiques, les francophones ont les meilleurs résultats dans la tâche de comptage syllabique. En revanche, les performances des anglophones sont supérieures à celles des francophones dans la tâche de suppression phonémique ainsi que dans toutes les tâches de jugement de similarité, excepté celles impliquant la syllabe et la rime. Les francophones semblent donc plus sensibles aux unités syllabiques que les anglophones qui manifestent une sensibilité accrue aux unités intra-syllabiques. Enfin, les anglophones ont obtenus des scores inférieurs à l'épreuve de lecture en 1ère année par rapport aux francophones. Ces données confirment l'idée que la consistance des relations graphophonologiques joue un rôle important dans l'apprentissage de la lecture et montrent l'impact négatif de l'opacité de l'orthographe. Les analyses de régression ont permis d'identifier les trois meilleurs prédicteurs du niveau ultérieur en lecture chez les deux groupes. Il s'agit, chez les anglophones, des capacités de manipulation des unités intra-syllabiques de type attaque-rime et chez les francophones, des capacités portant sur les unités syllabiques. Le troisième prédicteur, commun aux deux groupes, renvoie à la connaissance des lettres.

Mann & Wimmer (2002) ont examiné les prédicteurs de l'apprentissage de la lecture en anglais et en allemand. A la fin de la GSM, du grade 1 et du grade 2, les enfants ont passé deux tests de conscience phonologique, une tâche de dénomination de couleurs, une tâche d'identification de lettres et un test de lecture de mots et de non mots. Les analyses de régrssion montrent que la conscience phonologique est un prédicteur puissant de la lecture en anglais. En allemand, c'est la vitesse d'accès au lexique ou RAN qui est le meilleur et l'unique prédicteur de la lecture alors que la contribution de la conscience phonologique n'est pas significative.

Georgiou, Parilla & Papadopoulos (2008) ont réalisé une étude longitudinale dans le but de déterminer les prédicteurs de l'apprentissage de la lecture chez des enfants apprenant à lire une langue à orthographe opaque (anglais) et une langue à orthographe transparente (grec). L'étude a porté sur 110 enfants anglais et 70 enfants grecs. En 1ère année, des mesures de conscience phonologique, de mémoire phonologique, de dénomination rapide et des processus orthographiques ont été réalisées. Le niveau de lecture a été évalué chez les mêmes enfants en 2ème année. Les résultats de la modélisation par les équations structurelles indiquent que les connaissances phonologiques et orthographiques contribuent à la performance en lecture en 1ère et 2ème année. Cependant, l'importance de ces prédictions est différente dans les deux langues. En anglais, les meilleurs prédicteurs des performances d'identification de mots sont la conscience phonologique et les processus orthographiques. En grec, le meilleur prédicteur de la lecture est la dénomination rapide alors que les processus orthographiques ne contribuent pas significativement à la variance en lecture. Les auteurs concluent que les caractéristiques de la langue d'apprentissage est un facteur important qu'il est nécessaire de prendre en compte dans les études prédictives.

L'étude de Besse, Demont & Gombert (2007) avait pour objectif de déterminer dans quelle mesure les acquisitions de l'écrit en langue maternelle influencent les connaissances phonologiques et morphologiques (morphosyntaxiques et morphodérivationnelles) développées lors de l'apprentissage du français. Deux langues très différentes ont été comparées : la langue arabe qui présente un système graphique très différent du français et la langue portugaise qui, sous sa forme écrite, est la langue latine la plus proche du français. Différentes épreuves implicites et explicites évaluant le niveau de conscience phonologique et morphologique (flexionnelle et dérivationnelle), ainsi qu'un test de lecture de mots ont été administrées aux enfants. Les résultats indiquent que les arabophones ne sont pas sensibles à la rime mais seraient plus compétents à analyser explicitement la structure morpho-dérivationnelle des mots que les lusophones. De plus, la supériorité des arabophones à traiter les aspects morpho-dérivationnels par rapport aux lusophones est moins marquée dans les tâches implicites. Les arabophones seraient plus aptes à analyser consciemment la structure morpho-dérivationnelle de mots français car ce type d'analyse grapho-sémantique, déjà sollicité lors de la lecture en arabe, est libéré des caractéristiques de surface propres aux langues. Les résultats aux analyses corrélationnelles entre le décodage et les niveaux de connaissances phonémiques et morphologiques explicites montrent que, chez les arabophones, les connaissances liées au décodage, même très faiblement, semblent être

exclusivement les connaissances phonémiques. Cependant, à la différence des arabophones, les lusophones semblent investir leurs connaissances morphologiques explicites pour décoder en français mais pas leur conscience phonémique.

Ces études fournissent des résultats contradictoires en ce qui concerne la contribution des compétences phonologiques à l'apprentissage de la lecture dans une langue alphabétique qui varie selon la consistance de l'orthographe. Cependant, il est possible que les différences trouvées au niveau de la valeur prédictive des compétences phonologiques soient dues aux différentes mesures utilisées pour évaluer la performance en lecture : en anglais, les chercheurs mesurent essentiellement l'exactitude de la lecture (reading accuracy) tandis que dans les langues à orthographe consistante, ils évaluent surtout la vitesse de la lecture (reading speed).

1.3. Etudes comparatives entre bons lecteurs et faibles lecteurs

L'importance de la conscience phonologique est également démontrée à travers l'étude de cas d'enfants souffrant de troubles sévères d'apprentissages de la lecture. En effet, il a été montré que la majorité des enfants dyslexiques souffre de troubles phonologiques qui se manifestent par des difficultés importantes de répétition des non-mots et des difficultés à manipuler volontairement les sons qui composent les mots (Sprenger-Charolles & Colé, 2003). Il semble donc que des déficits touchant les procédures phonologiques soient à l'origine de troubles sévères d'apprentissage de la lecture. Les travaux comparant les bons et les faibles lecteurs corroborent cette position en montrant que les enfants faibles lecteurs présentent des performances nettement inférieures à celles des bons lecteurs, toutes unités de connaissance confondues. Ainsi, les résultats des expériences de Lundberg (1982) portant sur des sujets normolecteurs et faibles lecteurs de même âge et de même niveau intellectuel indiquent que les enfants du second groupe se différencient de ceux du premier par leurs capacités métaphonologiques moindres évaluées lors de tâches impliquant une analyse phonémique ou syllabique explicite. Morais, Gluytens & Alegria (1984) trouvent que les enfants dyslexiques âgés de 6 à 9 ans sont nettement moins performants que les lecteurs normaux dans une tâche de segmentation de mots ou de pseudo-mots en phonèmes alors que leurs performance dans une tâche équivalente portant soit sur des syllabes, soit sur un matériel non linguistique étaient de loin supérieures et assez proches de celles des enfants du groupes contrôle. Sabatier & al. (2003) constatent que les dyslexiques phonologiques présentent en fait un déficit de conscience linguistique global puisqu'ils sont dans l'incapacité d'accéder aux unités

minimales distinctives de la langue (les phonèmes et les morphèmes). Les sujets dyslexiques ont montré des difficultés lors des exercices de conscience phonologique explicite. Cependant, ils se sont montrés meilleurs dans les exercices de conscience morphologique. L'ensemble de ces études met en évidence que les dyslexiques ont des difficultés à la fois dans l'utilisation de l'information phonologique en lecture et dans la manipulation explicite ou consciente des unités phonologiques.

Par ailleurs, des travaux de recherche récents montrent que l'enfant dyslexique peut développer de façon " naturelle " une procédure compensatoire de lecture de type morphologique pour compenser, par exemple, la difficulté à recourir à la procédure phonologique (Casalis et Louis-Alexandre, 2000, Casalis, Colé & Royer, 2003). Les résultats témoignent que les performances des dyslexiques dans le domaine de la morphologie ne sont pas si faibles que celles obtenues dans le domaine de la phonologie. Casalis, Colé & Royer (2003) suggèrent que les enfants dyslexiques pourraient utiliser d'autres informations et s'appuyer sur d'autres processus pour résoudre ces tâches d'analyse morphologique que les enfants normolecteurs. L'étude a porté sur 12 enfants dyslexiques, 14 enfants de même âge chronologique et 15 enfants du même niveau en lecture que les enfants dyslexiques. La conscience morphologique a été évaluée à travers différentes tâches : choix de dérivés en contexte, recherche de l'intrus morphologique et production de formes dérivées. L'ensemble des résultats de cette recherche indique que les habiletés morphologiques évoluent au cours du développement et différencient bons et faibles lecteurs. Les dyslexiques ont réussi à développer une assez bonne conscience morphologique. Ils utilisent des stratégies qui s'appuient plus largement sur la signification, ce qui leur permettrait de compenser en partie leur déficit dans le traitement phonologique.

Ces résultats sont conformes à ceux observés chez des enfants en difficulté d'apprentissage de la lecture en arabe. Belajouza (2003) a cherché à explorer le fonctionnement de la reconnaissance des mots chez des enfants en grande difficulté d'apprentissage de la lecture. Les résultats révèlent que les élèves faibles lecteurs, de 5ème et 6ème années, présentent des déficits phonologiques et morphologiques importants. Ils obtiennent des scores équivalents à ceux obtenus par des normo-lecteurs de 2ème et 3ème années dans une tâche de suppression de phonèmes. En dépit de leur avancement dans la scolarité et une plus grande familiarité avec l'arabe standard, les faibles lecteurs n'ont pas progressé dans l'appropriation des connaissances phonologiques. Cette insuffisance dans le développement des connaissances phonologiques les empêche d'exploiter les règles de correspondance graphophonologique (la

voie d'assemblage) qui assurent le décodage des mots et surtout les pseudo-mots. Les faibles lecteurs obtiennent également un niveau très inférieur aux connaissances morphologiques. En effet, la reconnaissance des liens morphologiques semble s'appuyer plus sur l'identification globale du squelette consonantique que sur des indices phonétiques. Ammar (1997), de son côté, a trouvé que les faibles lecteurs de 1ère, 2ème et 3ème années présentent des scores inférieurs à ceux des bons lecteurs à des tâches de comptage, de segmentation et de suppression syllabique et phonémique. Dans la condition phonémique, les bons lecteurs et les faibles lecteurs réussissent les manipulations phonémiques des items en arabe standard ou dialectal de manière égale.

Abu-Rabia, Haifa, Share & Mansour (2003) ont examiné le rôle des processus phonologiques et orthographiques et des processus cognitifs dans l'identification des mots écrits en arabe. Trois groupes de lecteurs ont été comparés : des faibles lecteurs (9.5–10.5 ans), des normolecteurs de même âge chronologique et des normolecteurs plus jeunes de même niveau de lecture (7.5 – 8.5 ans). Les résultats révèlent des différences significatives entre les faibles lecteurs et les normolecteurs dans les différentes mesures de lecture et des mesures cognitives. La différence la plus forte concerne les compétences phonologiques. Ce retard en phonologie peut être lié au manque d'exposition à l'arabe littéraire à la maison. Les enfants faibles lecteurs obtiennent des scores très bas dans les différentes mesures phonologiques mais ils semblent plus performants en ce qui concerne les processus orthographiques, ce qui révèle leur stratégie de lecture basée sur les processus visuo-orthographiques. De même, il semble que le niveau de maîtrise des habiletés morphologiques, comme celui des habiletés phonologiques, soit corrélé avec le niveau des compétences en identification des mots. Ces données sont en accord avec les résultats obtenus en anglais. Par ailleurs, la mémoire visuelle, comme la mémoire de travail, apparaît comme un facteur important dans l'acquisition de la lecture en arabe contrairement à l'anglais (Ellis & Large, 1987). Cette différence peut être expliquée par le fait que la lecture en arabe implique l'utilisation des processus visuo-spatiaux pour identifier les lettres, les racines, les affixes et les voyelles courtes au dessus et au dessous des consonnes (Abu Rabia, 2001). Enfin, il semble que les faibles lecteurs arabophones présentent des caractéristiques similaires à celles reportées dans la littérature. En revanche, ils utilisent des stratégies visuo-orthographiques dès un âge avancé.

Dans une étude plus récente, Abu-Rabia (2007) avait pour objectifs d'explorer les compétences phonologiques et morphologiques chez des normolecteurs et des dyslexiques arabophones de différents âges et de mesurer la valeur prédictive de ces variables sur

l'identification de mots et la compréhension en lecture. Les résultats de cette étude montrent que les nomolecteurs ont des performances significativement meilleures que les dyslexiques dans toutes les mesures de lecture et d'écriture. Les dyslexiques montrent des difficultés dans le traitement des voyelles courtes qui indiquent la prononciation correcte et la fonction grammaticale des mots. Cette différence significative entre les deux groupes de lecteurs s'observe également dans les différentes mesures langagières. Ainsi, la dyslexie en arabe se caractérise par des symptômes similaires à ceux observés sur d'autres langues alphabétiques.

L'étude de Boumaraf & Badaoui (citée par Boukadida, 2008) a porté sur des enfants arabophones : des normo-lecteurs âgés entre 7 /8 ans et 9/10 ans et des dyslexiques âgés entre 9 et 10 ans. Les résultats montrent que les enfants dyslexiques éprouvent plus de difficultés à manipuler les petites unités (intrasyllabiques) et les syllabes fermées (CVC). De plus, Les enfants de 7/8 ans, tout comme les dyslexiques, n'arrivent pas à manipuler correctement les voyelles longues.

Boukadida & al. (2009) ont cherché à explorer le développement des connaissances phonologiques et morphologiques chez des apprentis lecteurs arabophones. Les enfants ont été distingués en bons (BL) et faibles lecteurs (FL) graçe à un test de lecture. Les données recueillies montrent que les BL développent une meilleure conscience phonologique que les FL. Ainsi, ils sont meilleurs dans la manipulation des unités phonémiques que les FL, en 1ère comme en 2ème année. Les résultats indiquent également que les BL se distinguent des FL au niveau des connaissances morphologiques. Les BL réussissent mieux la tâche de reconnaissance de l'appartenance à la même famille morphologique que la tâche de détection de l'intrus de rime.

1.4. Etudes avec entraînement métalinguistique

De nombreux travaux de recherche ont évalué les effets d'entraînements sur des enfants à risque de difficultés d'acquisition du langage écrit ou sur des enfants en situation d'échec en lecture (Lecoq, 1991 ; Zorman, 1999). L'objectif des différentes études est de comparer une situation expérimentale où l'enfant bénéficie d'un entraînement explicite quotidien et une situation ordinaire où l'enfant n'a pas d'autres aides que la pédagogie habituelle. Conformément aux nombreux travaux montrant les liens entre la conscience phonémique et l'apprentissage de la lecture, l'entraînement de type métaphonologique contribue au développement des processus de reconnaissance de mots écrits, centraux dans l'acquisition de la lecture (Ecalle, Magnan, Bouchafa, 2002). En effet, si les capacités métaphonologiques

sont causalement reliées à l'acquisition de la lecture, un entraînement dans ce domaine doit avoir un effet sur la lecture. Treiman (1991) avance que la capacité d'analyse phonémique est en relation directe avec les performances en lecture et qu'un entraînement de cette capacité améliore la capacité des enfants à apprendre les correspondances entre les formes écrites et parlées des mots. Cependant, les règles de transcription au niveau des phonèmes ne semblent ni faciles, ni naturelles au cours des phases précoces de l'apprentissage de la lecture (Bradley & Bryant, 1983). D'autres recherches ont montré que l'entraînement de la conscience phonologique consistant en une pratique régulière d'exercices oraux fait progresser plus rapidement l'enfant prélecteur et favorise l'apprentissage de la lecture (Lundberg, 1988 ; Vellutino, 1985 ; Wise & Olsson, 2004, cités par INSERM, 2007).

Lecoq (1991) a effectué une étude longitudinale comportant des programmes d'entraînement sur les capacités métaphonologiques, la mémoire de travail et la mémoire immédiate. Cinq groupes étaient constitués comportant entre 14 et 19 sujets. Trois de ces groupes étaient des groupes expérimentaux et recevaient un entraînement spécifique. Les deux autres étaient des groupes contrôles dont l'un a bénéficié d'un entraînement logico-mathématique. Les enfants ont suivis 12 séances d'entraînement, à raison de deux fois par semaine, pendant six semaines. Testés à la fin du CP, les sujets du groupe expérimental, entraînés aux habiletés métaphonologiques, ont obtenus les meilleurs résultats en lecture que ceux des deux autres groupes. L'entraînement de ces habiletés a donc une incidence significative sur le niveau atteint ultérieurement en lecture.

Ball & Blachman (1991) ont adopté également cette méthodologie en sélectionnant pour leur étude des enfants non lecteurs d'âge moyen 5 ans et 8 mois. Le premier groupe reçoit un entraînement phonémique (apprendre le lien existant entre les lettres de l'alphabet et les phonèmes, mettre ensemble les mots ayant la même sonorité, etc.), le second reçoit un enseignement basé sur le langage oral (écouter et comprendre des histoires, apprendre à catégoriser sémantiquement des images, etc.). Le groupe contrôle ne bénéficie d'aucun entraînement particulier. Les résultats indiquent que les enfants ayant bénéficiés d'un entraînement phonémique se différencient dans les épreuves de segmentation phonémique, d'orthographe, d'identification de mots et de lecture de mots réguliers. Ces travaux permettent donc de penser que l'acquisition de la lecture est fonction du niveau de développement métaphonologique. L'activation de cette performance préalablement à l'apprentissage lui-même aurait un effet facilitateur sur ce dernier.

L'étude de Zorman (1999), réalisée sur des élèves français de la Grande Section de Maternelle, montre l'efficacité de l'entraînement de la conscience phonologique sur l'apprentissage de la lecture. La méthode consiste à évaluer la conscience phonologique de 2265 enfants, puis à entraîner les sept enfants par classe qui ont le moins bon niveau de conscience phonologique pendant 31 jours à raison de 15-20 minutes par jour, puis à effectuer trois mois après une nouvelle mesure de la conscience phonologique et enfin à comparer les progrès réalisés entre les deux périodes. Les résultats mettent en évidence les différences de score de conscience phonologique avant et après l'entraînement. Les élèves, quel que soit leur niveau de départ aux épreuves phonologiques, ont plus progressé dans le groupe entraîné que dans le groupe contrôle. Ainsi, un entraînement phonologique en petit groupe des enfants les plus faibles en capacités métaphonologiques leur permettait de faire des progrès significativement supérieurs à ceux qui n'en bénéficiaient pas. Selon l'auteur, ce type d'entraînement devrait aussi être réalisé en même temps que l'apprentissage de la lecture en CP et au-delà pour les élèves qui nécessitent la mise en place de remédiations pédagogiques.

L'étude de Lyster (2002) montre l'effet bénéfique d'un entraînement morphologique, notamment dans sa composante sémantique, sur les capacités d'identification de mots. Ainsi, les enfants ayant suivi un entraînement morphologique en grande section maternelle, montrent six mois plus tard des améliorations en lecture en comparaison avec un groupe contrôle. Ainsi, les capacités métamorphologiques seraient liées, comme les capacités métaphonologiques, à l'apprentissage de la lecture d'une langue alphabétique mais ne seraient pas réinvesties exclusivement dans l'activité de décodage.

Casalis & Colé (2009) ont comparé les effets d'un entraînement phonologique et d'un entrainement morphologique chez des enfants en GS. Les résultats montrent une influence du travail sur la morphologie sur le développement des habiletés phonologiques. Ainsi, l'entraînement morphologique augmente la sensibilité des enfants aux unités phonologiques sans pour autant favoriser la manipulation explicite des phonèmes. De même, l'entraînement phonologique aide les enfants à segmenter les mots en leurs constituants morphémiques mais non à produire des mots dérivés.

2. Prédiction de la réussite scolaire fournie par les enseignants

Les enseignants, en contact quotidien avec les élèves, sont eux aussi en mesure d'observer avec précision le fonctionnement de leurs élèves à l'école et de repérer certaines difficultés. Leurs évaluations ou appréciations sont corrélées avec des évaluations de type psychologique réalisées dans la même période et s'avèrent prédictifs de la réussite scolaire ultérieure (Ecalle, 2000a, 2000b ; Florin, 1991 ; Florin & al., 1999 ; Guimard & Florin, 2007). Plusieurs études indiquent en outre que les évaluations par les enseignants en fin de maternelle ou en début de 1ère année primaire prédisent mieux les performances académiques ultérieures ou le parcours scolaire des élèves que les évaluations cognitives (Ecalle, 2000a, Quay & Steel, 1998). En s'intéressant aux rapports entre le comportement verbal des élèves en maternelle et leur réussite scolaire ultérieure, Florin (1991) a observé que le comportement psychologique de l'élève en classe évalué par l'enseignant constitue un prédicteur puissant. Ce comportement est défini en particulier par le niveau de langage, le niveau d'attention et la capacité à suivre le rythme de travail en classe. Il s'avère que les difficultés scolaires ultérieures sont prédictibles pour les élèves qui, dès 4-5 ans, sont décrits avec un faible niveau de langage, des capacités d'attention réduites et des difficultés à assurer les tâches scolaires sans être contraints ou soutenus.

Dans une étude longitudinale, Guimard & Florin (2001) ont étudié la valeur prédictive d'une évaluation des comportements scolaires par des enseignants en moyenne section de maternelle (MSM) sur les comportements scolaires (principalement en lecture) en CP et en CE2. Deux types d'évaluation ont été réalisées en MSM : une évaluation des comportements scolaires effectuée par les enseignants pour chaque enfants à partir d'un questionnaire (11 items) et des évaluations des compétences cognitives et langagières réalisées par des psychologues (processus séquentiels et simultanés du K-ABC, compréhension orale, morpho-syntaxe et compréhension de concepts spatio-temporels). Les performances scolaires prises en compte aux différents niveaux scolaires (CP et CE2) sont principalement des compétences en lecture évaluées par deux épreuves, lecture et déchiffrement du K-ABC pour le CP et identification de mots pour le CE2. Trois autres indicateurs de compétences scolaires ont été considérés en CE2 : les résultats à une épreuve de lecture-compréhension d'énoncés et les performances des enfants aux évaluations nationales de mathématiques. Les analyses de corrélation révèlent que les évaluations comportementales réalisées en maternelle sont dans l'ensemble fortement liées

au niveau de lecture en CP et en CE2. Toutefois, les liaisons entre les variables sont moins nombreuses en CE2 qu'en CP. Avec l'avancement dans la scolarité, certaines variables semblent perdre de leur pouvoir alors que d'autres semblent en gagner. C'est particulièrement le cas de la « confiance en soi » et de la « capacité à organiser son travail » qui apparaissent comme les meilleurs prédicteurs des résultats aux diverses évaluations scolaires réalisées en CE2. Les comportements scolaires présentent une forte stabilité prédictive dans le temps puisque, en CP comme en CE2, ils contribuent pour une part modeste mais non négligeable et constante, à l'explication des performances en lecture.

Florin, Khomsi, Guimard, Ecalle & Guéguan (1999) ont tenté de montrer que les évaluations des enseignants sur la seule participation des élèves aux activités langagières en grande section de maternelle permettent de distinguer des groupes aux compétences langagières et cognitives contrastées. Trois niveaux ont été relevés, les grands, les moyens et les faibles parleurs (GP, MP et FP). Au début de GSM, des épreuves d'identification de mots à l'oral, de compréhension du langage oral, du jugement de grammaticalité et de logique non verbale ont été proposées aux élèves. Au début de CP, une épreuve d'évaluation du niveau de conceptualisation de la langue orale et écrite a été également administrée aux enfants. Les résultats montrent que les performances décroissent des grands parleurs, aux moyens et aux faibles parleurs. L'effet groupe est significatif pour les performances en identification de mots à l'oral et en logique non verbale. En début de CP, cet effet persiste pour l'épreuve de conceptualisation de la langue orale et écrite. En relevant le niveau de participation verbale des élèves aux activités langagières précoces dès le début de la GSM, les enseignants sont capables de distinguer les élèves aux performances langagières et cognitives plus ou moins contrastées.

Les résultats de l'étude longitudinale de Guimard, Cosnefroy & Florin (2007) portant sur un panel de 5549 élèves suivis de la 1ère année primaire à l'entrée en 6ème de collège mettent en évidence qu'une évaluation par un questionnaire effectuée par les enseignants en début de CP explique significativement les performances en français et en mathématique et le parcours scolaire des élèves. Les analyses de régression montrent que les évaluations scolaires standardisées réalisées en 1ère année sont les meilleurs prédicteurs des performances scolaires mesurées en 3ème et 6ème année ; mais l'évaluation des comportements scolaires réalisée par les enseignants ajoute une part significative non négligeable à l'explication des performances académiques des élèves. L'évaluation des enseignants apporte donc des informations

différentes et complémentaires de celles obtenues au moyen des évaluations standardisées (Guimard & Florin, 2008).

Ecalle (2000b) a cherché à dégager les prédicteurs de réussite en identification de mots écrits et en compréhension en lecture. Pour atteindre cet objectif, 68 enfants ont été suivis longitudinalement de la GSM au CE1. Les évaluations des enseignants ont été associées aux mesures directes prélevées auprès des enfants. A chaque passage dans les classes, il était demandé aux enseignants d'évaluer les élèves à l'aide d'une échelle en 7 points, dans trois domaines : langage, attention et capacité à suivre le rythme de travail en classe (LAT). Les résultats montrent que l'indicateur LAT en début de CP intervient en tant que premier prédicteur pour expliquer les performances en orthographe lexicale (6 mois après) et les performances en compréhension en lecture (18 mois après) et en second prédicteur pour expliquer les performances en identification de mots en fin de CE1. Ces données montrent que les appréciations des enseignants sont parmi les plus prédictives des performances ultérieures des élèves tant en lecture qu'en orthographe. Ce jugement, fourni dès le début de CP, s'avère suffisamment précis pour prédire la réussite scolaire ultérieure dans les trois domaines relatifs à l'apprentissage de la langue écrite : identification et production de mots et compréhension en lecture.

Meljac, Kugler & Mogenet (2001) ont mené une étude longitudinale pour tenter de savoir dans quelles mesures les enseignants peuvent prédire la réussite et l'échec dans l'apprentissage de la lecture. Avec le concours des enseignants de maternelle, ils ont constitué au sein de plusieurs écoles situées dans des zones de niveau socio-économique moyen, deux groupes d'enfants : l'un pour lequel les enseignants émettaient, quant à l'apprentissage de la lecture, un pronostic favorable ; l'autre pour lequel le pronostic était défavorable. Les enfants sont suivis de la GSM au CP. Le niveau de lecture a été évalué en fin de CP par deux épreuves de lecture, l'une consistait à lire un texte court et à en restituer quelques éléments, l'autre à identifier des mots isolés de complexité croissante (épreuve extraite du K-ABC). Les résultats montrent que les prédictions des enseignants de maternelle se révèlent de bons indicateurs de l'évolution future d'un écolier au cours de la 1ère année du primaire, en particulier en ce qui concerne son acquisition de lecture. Les prédictions des enseignants ont été vérifiées dans 75% des cas.

3. Prédictions individuelles

Les études précédentes permettent de déterminer les compétences précoces qui ont une incidence sur le futur niveau de lecture des enfants. Les évaluations prédictives globales restent cependant insuffisantes lorsqu'il s'agit de prédire le devenir des élèves pris individuellement. Diverses recherches (Guimard & Florin, 2007 ; Piquard-Kipffer, 2003 ; Peterson & Elbro, 1999) ont essayé de mesurer la valeur prédictive individuelle d'épreuves standardisées dans une perspective de prévention des difficultés d'apprentissage. La méthodologie utilisée consiste à évaluer les propriétés des outils : leur sensibilité, leur spécificité, leur valeur prédictive positive et leur valeur prédictive négative[33]. De façon générale, les résultats de ces recherches montrent que le pourcentage de prédictions correctes au niveau des individus est élevé (plus de 80%). Toutefois, des différences entre sensibilité et spécificité sont observées : alors que peu d'enfants diagnostiqués comme n'étant pas « à risque » ont en fait des difficultés de lecture, une partie de ceux prédits « à risque » apprend normalement à lire.

L'étude de Peterson & Elbro (1999) a porté sur 150 enfants danois. Six mesures prédictives ont été relevées avant l'apprentissage de la lecture : connaissance des lettres, compétences phonologiques et morphologiques (3 mesures), clarté de la prononciation et les antécédents de dyslexie dans la famille. Les résultats montrent que le devenir en lecture est prédit correctement pour 129 des 150 enfants (86%). La prédiction est toutefois moins fiable pour les enfants qui, effectivement, deviennent dyslexiques (60%) que pour ceux qui n'auront pas de difficulté en lecture (94%). L'étude de Piquard-Kipffer (2003) a porté sur 85 enfants français suivis de la grande section à la 2ème année primaire. A 8 ans, les enfants ont été séparés en deux groupes en fonction de leur niveau de lecture : les faibles lecteurs qui ont plus de 6 mois de retard en lecture et, les autres, qui ont moins de 5 mois de retard. Les prédicteurs sélectionnés pour classer les enfants sont ceux qui expliquent la majeure partie de la variance en lecture : capacités de discrimination phonémique et de dénomination sérielle rapide. A elles seules, ces deux épreuves permettent de classer correctement 83,5% des enfants en fonction de leur niveau de lecture en fin de la 2ème année primaire (71 sur 85). La prédiction

[33] La sensibilité d'un outil prédictif est déterminée par le pourcentage de sujets prédits en difficulté au temps T1 par rapport au nombre de sujets réellement en difficulté à T+n. Sa spécificité désigne le pourcentage de sujets prédits sans difficulté à T1 parmi les sujets ne présentant pas de difficulté à T+n. Sa valeur prédictive positive renvoie au pourcentage de sujets réellement en difficulté à T+n parmi ceux prédits en difficulté à l'évaluation initiale. Sa valeur prédictive négative désigne le pourcentage de sujets ne présentant réellement pas de difficulté à T+n parmi ceux prédits sans difficulté à l'évaluation initiale (T1).

est toutefois un peu moins fiable pour les enfants qui deviennent faibles lecteurs (76%) que pour ceux de l'autre groupe (87%). L'étude de Guimard & Florin (2007) vise à déterminer la validité prédictive globale et individuelle d'un questionnaire d'évaluation des comportements et des compétences scolaires renseigné par les enseignants chez des élèves suivis de la grande section de maternelle (GSM) au cours préparatoire (CP). D'autre part, l'objectif est de comparer ces prédictions à celles issues des évaluations cognitives et langagières standardisées. Les résultats indiquent que la valeur prédictive globale du questionnaire est bonne et supérieure à celle de la batterie cognitive et langagière. Toutefois, l'étude de la validité prédictive individuelle montre que les deux évaluations de GSM parviennent à mieux prédire les futurs bons lecteurs de CP (spécificité et valeur prédictive négative généralement supérieures à 90%) que ceux qui éprouvent des difficultés dans ce domaine (sensibilité et valeur prédictive positive inférieures ou égales à 50%). Les résultats de cette étude sont ainsi conformes à ceux obtenus par Peterson & Elbro (1999) et par Piquard-Kipffer (2003).

En définitive, l'ensemble de ces recherches soulignent les limites des outils d'évaluation utilisés pour prédire les difficultés ultérieures en lecture dans une population d'enfants « tout-venant ».

4. Apports et limites des études prédictives

Les évaluations cognitives et langagières standardisées sont largement utilisées pour décrire les difficultés d'apprentissage de la lecture et pronostiquer précocement les enfants à risque pour cet apprentissage. Le niveau de lecture a été prédit à partir d'habiletés diverses examinées à différentes époques préscolaires, les prédicteurs n'étant pas systématiquement les mêmes aux différentes périodes. Ainsi, le pouvoir prédictif des compétences phonologiques baisse en fonction du niveau scolaire alors que celui des capacités de dénomination rapide augmente (Kirby & al., 2003). De façon générale, la conscience phonologique se réinvestit préférentiellement au niveau des capacités de reconnaissance de mots, essentiellement au début de l'apprentissage, et s'avère déterminante en facilitant la compréhension du principe alphabétique. En revanche, elle ne s'avère pas directement liée aux capacités en lecture-compréhension. D'autres habiletés semblent reliées préférentiellement à la compréhension en lecture, notamment les connaissances morphologiques et syntaxiques (Muter & al., 2004) et la mémoire verbale (Demont, 2003). Ces résultats indiquent que les habiletés sous-tendant ces deux compétences ne sont pas de même nature.

Un premier type de problème des recherches longitudinales prédictives qui ont examiné le lien entre les habiletés cognitives et langagières et le développement ultérieur de la lecture vient du fait que certaines compétences sont très peu développées avant un certain âge. Par exemple, si le rôle de la conscience phonologique est bien connu, son évaluation est très difficile avant l'âge de 5 ans, compte tenu de la difficulté des tâches. Il ne peut donc constituer un prédicteur valide en deçà de cet âge. Ainsi, l'étude de Lecoq (1991) a mis en évidence que les capacités d'analyse phonologique se développent très lentement à certaines époques, et plus rapidement par la suite. A 4 ans, en effet, seules deux parmi les dix épreuves qui impliquent des unités phonémiques sont à peu près maitrisées, mais les enfants font des progrès substantiels dans ce domaine avant l'apprentissage de la lecture. A la différence des épreuves d'analyse phonologique, le développement des capacités de mémoire entre 3 ans et demi et 6 ans est peu important. De plus, les disparités entre épreuves pour les différentes tranches d'âge sont minimes. Enfin, pour les épreuves de dénomination rapide, en dehors de celles qui impliquent des chiffres, les performances n'évoluent pas notablement alors que, pour la vitesse d'articulation, une forte progression est relevée entre 5 ans et 5 ans et demi. Toutefois, ce ne sont pas nécessairement les tâches les mieux maitrisées à un moment donné qui sont les plus prédictives. En raison de cette progression non linéaire, des déficits dans certaines habiletés peuvent ne pas être détectables à certaines époques parce qu'ils sont peu développées avant un certain âge. Les prédicteurs peuvent également varier en fonction du moment où la mesure de lecture est prise.

Un autre type de problème concerne la pluralité des variables prédictives. Bien que la majorité des études s'attache à examiner un seul des prédicteurs potentiels, en l'occurrence les connaissances phonologiques, d'autres compétences sont associées à la lecture. Récemment, plusieurs études longitudinales ont rapporté que la capacité de mémoire verbale mesurée avant l'apprentissage formel de la lecture permet de prédire la performance en lecture un an plus tard (Demont, 2003 ; Rohl & Pratt, 1995). Cette relation s'explique par le fait qu'un codage efficace en mémoire de travail permettrait au lecteur débutant de maintenir une représentation exacte des phonèmes associés aux lettres ou groupes de lettres. D'un autre côté, il est clair qu'on ne peut conclure à l'existence d'un lien entre un prédicteur et la maîtrise de la lecture que dans la mesure où les autres prédicteurs potentiels sont contrôlés. En effet, la composante mnésique apparaît fortement impliquée dans les épreuves de conscience phonologique. Par exemple, la plupart des tâches phonologiques font appel, de façon plus ou moins importante, à la mémoire (par exemple juger si deux mots riment ou non). De même,

certaines tâches de mémoire font appel à une certaine forme de conscience segmentale (par exemple, la répétition de pseudo-mots). On ne pourra donc conclure à l'existence d'un lien spécifique entre habiletés phonologiques (ou mémoire) et lecture qu'en ayant contrôlé au préalable les capacités de mémoire (ou phonologiques). Il peut donc être postulé qu'une partie des déficits observés aux épreuves de conscience phonologique pourrait refléter tout autant un déficit mnésique que des difficultés d'analyse explicite de la structure phonologique de la parole. Il s'avère, en conséquence, déterminant d'étudier les liens entre ces deux habiletés en vue d'appréhender leurs rôles respectifs dans l'acquisition de la lecture et ses difficultés. C'est bien ce que signale l'étude de Rohl & Pratt (1995) qui montre que, lorsque les variables de conscience phonologique sont contrôlées, la mémoire de travail (répétition de mots à l'envers), mesurées au début de la 1ère année ne contribue pas de façon significative à l'acquisition de la lecture en 2ème année.

Certaines recherches soulignent les limites des batteries cognitives et langagières pour repérer les enfants en difficulté et prédire leur devenir (Zorman & Jacquier-Roux, 2004). Elles reprochent à certaines batteries leurs qualités psychométriques insuffisantes, l'absence d'étalonnage solide ou l'absence de validité prédictive. Enfin, certaines variables, comme par exemple les variables conatifs, n'ont quasiment jamais été prises en compte dans les études sur les facteurs prédictifs de la lecture. En effet, certains auteurs (Leybeart & al., 1994 ; Content & Leybeart, 1992) suggèrent que les enfants peuvent être exposés à des conditions d'apprentissage très variées au cours de la 1ère année qui auront des répercussions tant au niveau du résultat atteint en lecture et en orthographe que des mécanismes de traitement de l'information écrite et orale.

5. Conclusion

Les études présentées dans ce chapitre indiquent qu'il est possible de prédire de façon très fiable les futures difficultés de lecture et signalent les compétences précoces qui ont une incidence sur le niveau ultérieur de lecture des enfants. Les prédictions peuvent être globales ou individuelles, porter sur des enfants tout venants ou en difficulté, et être réalisées à plus ou moins long terme. Des études longitudinales ont montré la valeur prédictive du niveau de conscience phonologique des enfants avant l'apprentissage de la lecture, sur la réussite de celui-ci (Abu-Rabia, 2007 ; Gombert & Demont, 1996 ; Lecoq, 1991 ; Ibrahim & al., 2007 ; Muter & al., 2004 ; Schatschneider & al., 2004 ; Storch & Whitehurst, 2002). D'autres recherches ont mis en évidence qu'un entraînement de la conscience phonologique, avant

d'apprendre à lire, permet d'améliorer les performances en lecture (Ball & Blachman, 1991 ; Lecoq, 1991 ; Zorman, 1999). Un fort consensus se dégage sur le fait que les prédicteurs les plus puissants de l'apprentissage de la lecture sont, d'une part, les compétences en analyse phonémique, d'autre part, le niveau de connaissance des lettres et, enfin, les capacités de dénomination rapide. Les capacités mnésiques apportent également une part de variance explicative du niveau de lecture (Demont, 2003 ; Rohl & Pratt, 1995). En comparaison, le poids des habiletés non verbales, tout comme celui des facteurs socioculturels des enfants sont moins fortement reliés à leur futur niveau de lecture que ne le sont leurs habiletés linguistiques. Les enseignants jouent aussi un rôle important dans le repérage des difficultés des élèves et dans la mise en œuvre d'actions de prévention dans la mesure où leur contact quotidien avec les élèves leur permet de repérer des difficultés dans les différentes composantes de l'adaptation scolaire (Florin, 1991 ; Guimard & Florin, 2007 ; Guimard & Florin, 2001).

De façon générale, les données de la littérature montrent que les comportements des élèves en classe décrits par les enseignants et leurs évaluations des performances scolaires constituent des indicateurs particulièrement pertinents pour différencier les élèves et prédire leur réussite ou échec dans les apprentissages scolaires. En revanche, les évaluations prédictives globales restent insuffisantes lorsqu'il s'agit de prédire le devenir des élèves pris individuellement. De façon générale, les résultats des recherches montrent que la valeur prédictive individuelle des évaluations standardisées est acceptable (Piquard-Kipffer, 2003 ; Peterson & Elbro, 1999). La prédiction est toutefois moins fiable lorsqu'elle concerne les futurs faibles lecteurs que les futurs bons lecteurs.

En effet, il est primordial de détecter le plus rapidement possible les futurs faibles lecteurs afin de construire des projets d'aide adapté. Ainsi, des différences peu importantes chez le pré-lecteur ou le jeune apprenti lecteur peuvent se transformer en différences extrêmement importantes un an ou deux ans plus tard, ce qui renvoie à l'effet Matthieu (un bon lecteur deviendra encore meilleur, le faible lecteur encore plus faible)[34].

[34] Pour une revue, voir Stanovich (1986)

Deuxième partie :
Partie expérimentale

Chapitre 1: Perspectives de la recherche

1. Présentation de la problématique

L'apprentissage de la langue écrite constitue une des priorités de l'école. Les difficultés que l'élève rencontre dans ce domaine exercent de fortes contraintes sur son cursus scolaire et risquent de le placer progressivement en situation d'échec. En Tunisie, l'échec scolaire, qui touche entre 20 et 25% des enfants, a un coût économique et social élevé pour la communauté nationale (Ministère de l'éducation et de la formation, 2007). Les enjeux de la maîtrise de la lecture justifient l'importance que prend son apprentissage dès la première année du primaire et l'on comprend bien pourquoi, depuis si longtemps, on s'attache à identifier les facteurs reliés à sa réussite. De nombreux travaux soulignent l'intérêt d'une prévention précoce des difficultés scolaires des élèves (Guimard & Florin, 2008 ; Zorman & al., 2004 ; Zazzo, 1978). Il est donc primordial de chercher à identifier les déterminants de la réussite ou de l'échec pour pouvoir développer des moyens d'intervention et de remédiation adéquats. La présente recherche se veut être une contribution dans ce domaine.

Actuellement, grâce à de nombreuses études prédictives, les facteurs déterminants dans l'apprentissage de l'écrit sont bien connus (*cf.* partie théorique, chapitre 3 et 4). Plusieurs études longitudinales, effectuées sur des langues différentes, ont démontré que les habiletés précoces dans le domaine du langage oral peuvent être fortement et significativement corrélées avec les performances ultérieures en lecture (Storch & Whitehurst, 2002 ; Schatschneider & al., 2004). En effet, certaines de ces habiletés ne sont pas simplement associées aux capacités de lecture mais pourraient bien les déterminer. Ainsi, le niveau de conscience phonologique à l'école maternelle est un bon prédicteur de la réussite future dans l'apprentissage de la lecture (Demont & Gombert, 1996 ; Kirby & al., 2003 ; Lecoq, 1991 ; Schatschneider & al., 2004 ; Muter & al., 2004). Le statut des différentes unités phonologiques en tant que prédicteurs du développement ultérieur des compétences en identification de mots écrits a été explicité, entre autres, dans l'étude de Muter & al. (2004). En comparant le rôle des compétences phonémiques et celles portant sur la rime, ces auteurs ont constaté que le développement des capacités de reconnaissance de mots au cours des deux premières années de scolarité, est prédit principalement par différentes compétences de manipulation phonémique (inversion et suppression phonémiques). Par contre, la sensibilité à

la rime n'apparaît pas comme une unité importante pour l'identification de mots et ne semble pas prédire de façon significative le niveau ultérieur en lecture. Les études menées sur la langue arabe vont dans le même sens en démontrant la non pertinence de la rime dans l'apprentissage de la lecture (Ammar, 2002 ; 2003 ; Boukadida, 2008 ; Besse & al., 2007).

En arabe, la relation entre les différentes unités phonologiques et la lecture n'a pas été très étudiée. Les recherches menées sur les compétences phonologiques confirment le rôle primordial de la conscience phonémique lors de la reconnaissance des mots écrits. Ammar (1997) a trouvé que les élèves bons lecteurs de 1ère, 2ème et 3ème années réussissent moins bien les tâches de segmentation phonémique que leurs homologues anglais ou français. Il explique les difficultés rencontrées par les lecteurs arabes par la situation de diglossie puisque les mots changent principalement de représentations phonologiques lors du passage de l'arabe standard à l'arabe dialectal. Une étude de Ez-Zaher (2004, cité par Besse, 2007) confirme l'importance de la conscience phonémique en arabe au début de l'apprentissage de la lecture. La suppression de phonème apparaît être le meilleur prédicteur des habiletés de lecture une année plus tard, mais les compétences en segmentation syllabique apportent également une contribution non négligeable à l'explication de la variance en lecture. L'auteur explique ce résultat par le fait que les unités graphiques en arabe correspondent le plus souvent à la syllabe dû au marquage diacritique des voyelles, ce qui contraint le développement des capacités métaphonémiques et favorise, à l'écrit, une analyse en syllabes. Afin de vérifier ces résultats et déterminer le poids des différentes unités phonologiques dans l'apprentissage de la lecture en arabe, nous avons utilisé différentes mesures de sensibilité phonologique portant sur différentes unités (rime, syllabe, phonème) et sur différentes tâches (comptage, segmentation, suppression, détection de l'intrus).

De même, plusieurs recherches récentes ont montré l'importance de la connaissance des lettres, considérée comme l'un des plus puissants prédicteurs préscolaires de l'apprentissage de la lecture (Biot-Chevrier & al., 2008 ; Ecalle & al., 2009 ; Muter & al., 1998 ; Piquart-Kipffer, 2003 ; Schatsneider & al., 2004 ; Scarborough, 1998). Ces travaux ont démontré que la connaissance du nom des lettres contribue à l'acquisition de la lecture selon plusieurs voies. La connaissance du nom des lettres permet aux prélecteurs de relier les mots oraux et les mots écrits au moyen d'associations sublexicales entre lettres et noms de lettres (Foulin, 2005 ; 2007 ; Foulin & Pacton, 2006). Le nom des lettres intervient également dans le développement initial de l'identification des mots écrits. Les prélecteurs qui connaissent le nom des lettres ne se limitent pas à identifier les mots écrits à partir d'indices visuels mais

sont capables d'utiliser l'information phonologique lorsque des noms des lettres sont présents dans la prononciation des mots. La connaissance du nom des lettres est aussi un facteur de progrès dans l'apprentissage du son des lettres et dans le développement de la sensibilité phonémique (Ecalle & al., 2009 ; Foulin & Pacton, 2006). Des recherches récentes ont montré que la connaissance du nom des lettres joue un rôle primordial dans l'acquisition de l'écrit au cours de la période préscolaire et scolaire (Foulin, 2007) et constitue un puissant prédicteur de l'apprentissage de la lecture (Biot-Chevrier & al., 2008 ; Ecalle & al., 2009). Ainsi, Muter & al. (1998) trouvent que la connaissance des lettres et l'habileté de segmentation phonémique mesurées au début de l'apprentissage sont des prédicteurs puissants de la réussite de la lecture une année plus tard. Les relations prédictives entre l'identification des lettres et la capacité de lecture indiquent que les enfants qui débutent dans l'apprentissage formel de la lecture avec une plus forte maîtrise des lettres apprennent à lire plus vite et mieux que les autres (Foulin & Pacton, 2006). Bond & Dykstra (cité par Muter & al., 2004) montrent que la connaissance des lettres au début de la 1ère année prédit 26 à 36 % de la variance des performances en identification de mots en fin de l'année.

Les travaux de recherches présentés dans les différents chapitres théoriques montrent que d'autres connaissances métalinguistiques pourraient aussi influencer le développement des procédures de reconnaissance de mots écrits telles que les capacités d'analyse morphologique ou conscience morphologique qui renvoient aux connaissances morphologiques que l'enfant sait utiliser et manipuler de façon consciente. Les études portant sur le développement des capacités métamorphologiques ont montré que celles-ci peuvent intervenir de façon importante et précoce dans l'apprentissage de la lecture (Colé & al., 2003 ; Marec-Breton, 2009 ; 2003 ; Marec-Breton & al., 2005). Finalement, il a aussi été mis en évidence un lien entre les performances en lecture et les habiletés des enfants à accéder facilement et rapidement à l'information phonologique en mémoire à long terme. Cette habileté a été évaluée avec des tâches de dénomination où on demande à l'enfant de dénommer, le plus vite possible, une série d'items (des chiffres, des couleurs, des objets) représentés graphiquement. Les différences de vitesse avec laquelle les enfants pré-lecteurs sont capables d'accomplir cette tâche prédisent fortement les différences ultérieures dans l'acquisition des habiletés d'identification de mots (Castel & al., 2008 ; Kirby & al., 2003).

Un large consensus semble ainsi se dégager sur le fait que les compétences précoces les plus fortement reliées aux performances ultérieures en lecture sont celles qui sont directement impliquées dans cette activité, comme par exemple la connaissance des lettres et les capacités

d'analyse phonologique. Actuellement, des études de plus en plus nombreuses amènent à tenir compte du rôle de certaines compétences cognitives dans le développement des performances en lecture. Les résultats d'études longitudinales mettent en évidence le pouvoir prédictif de la mémoire du travail (Lecoq, 1991 ; Rohl & Pratt, 1995) et du traitement visuo-attentionnel (Bosse & valdois, 2009 ; Bosse & al., 2009) évalués en maternelle sur l'efficience ultérieure en lecture. Une hypothèse récente accorde un rôle primordial au contrôle exécutif dans l'apprentissage de la lecture (Boulc'h, 2008 ; Boulc'h & al., 2007).

Les résultats disponibles sur l'arabe et l'hébreu confortent les principes généraux des modèles d'acquisition de la lecture dans les langues à écriture latine mais révèlent aussi des différences liées aux spécificités linguistiques des langues sémitiques, notamment en mettant en évidence un rôle plus important de la morphologie (Abu-Rabia, 2007 ; Abu-Rabia & al., 2003 ; Boukadida, 2008 ; Ben Dror & al., 1995 ; Grainger & al., 2003). L'arabe est une langue alphabétique qui présente des propriétés linguistiques particulières. En arabe, le mot a une structure bidimensionnelle. Il est fondé sur une racine purement consonantique, les voyelles ne sont que des éléments de dérivation. C'est à partir de la racine que la plupart des mots sont formés par affixation ou par alternance vocalique. L'arabe est composé de deux systèmes d'écriture, l'un vocalisé destiné surtout aux apprentis lecteurs ; l'autre non vocalisé constitue l'écriture standard. Ce double système d'écriture renforce la transparence de la structure bidimensionnelle du mot. S'interroger sur le rôle des connaissances morphologiques dans l'apprentissage de la lecture en langue arabe s'avère donc parfaitement légitime. Les caractéristiques orthographiques et morphologiques de l'arabe font que l'analyse morphologique est nécessaire lors de la reconnaissance des mots écrits, et qu'elle joue un rôle plus important que dans les langues latines (Gombert, 2003).

La diglossie représente une réalité linguistique pour les apprentis lecteurs arabes. En Tunisie, on est en présence de deux variétés d'une même langue : l'arabe standard, niveau écrit et lu, qui est la langue des enseignements scolaires et l'arabe dialectal, niveau essentiellement parlé qui représente la langue maternelle pratiquée par tous les tunisiens avec certes, des variantes régionales au niveau phonologique et lexical. La distance, notamment phonologique, entre l'arabe dialectal et standard est souvent perçue comme une contrainte pour l'apprenti lecteur arabophone et serait à l'origine de nombreuses difficultés d'apprentissage de la lecture (Belajouza & Snoussi, 2006 ; Derguini & Mekidech, 2003 ; Feitelson & al., 1993 ; Saiegh-Haddad, 2004).

Les travaux de recherche réalisées sur l'apprentissage de la lecture en contexte tunisien ont essayé d'étudier l'évolution des stratégies d'identification de mots écrits en arabe (Ammar, 1997) et le rôle des connaissances phonologiques et morphologiques dans cet apprentissage (Boukadida & al., 2009 ; Boukadida, 2008 ; Besse, 2008) et d'effectuer une comparaison interlangue entre l'arabe et une autre langue alphabétique (Besse, 2008 ; Besse & al., 2007). Dans la continuité de ces rechreches, nous avons essayé dans cette étude d'une part de confirmer certains résultats déjà obtenus sur les langues à écriture latine, comme en ce qui concerne l'importance des connaissances phonologiques et, d'autre part, de révéler des différences liées surtout aux spécificités de la langue arabe et du contexte d'apprentissage caractérisé par la situation de diglossie. Nous avons pensé également à l'intérêt qu'il y aurait à découvrir le rôle d'autres facteurs langagiers comme la connaissance des lettres et la rapidité d'accès au lexique et des facteurs cognitifs comme la mémoire à court terme et le niveau attentionnel. Le rôle de ces facteurs dans l'apprentissage de la lecture en langue arabe n'a pas été à notre connaissance étudié.

La présente recherche vise précisément à déterminer les facteurs précoces contribuant à l'apprentissage de la lecture en langue arabe. Nous espérons ainsi dégager les meilleurs prédicteurs du niveau ultérieur en lecture. La recherche des facteurs qui déterminent la maîtrise de la lecture ne peut se faire efficacement qu'en adoptant une approche longitudinale en suivant depuis la classe préparatoire jusqu'à la fin de la 2ème année primaire un groupe d'enfants tunisiens afin d'évaluer le poids prédictif de certaines compétences langagières et cognitives sur leur niveau ultérieur de lecture.

Les évaluations cognitives et langagières standardisées sont largement utilisées pour prédire la réussite et les difficultés d'apprentissage de la lecture et repérer précocement les enfants à risque pour cet apprentissage (Guimard & Florin, 2008). Toutefois, pour mieux comprendre les difficultés scolaires des élèves, il peut s'avérer utile de ne pas limiter l'évaluation aux seuls domaines cognitif et langagier, mais de disposer également d'une évaluation des comportements et des compétences des élèves en classe réalisées par les enseignants. Plusieurs études (Ecalle, 2000a ; Quay & Steele, 1998) indiquent que les évaluations des comportements scolaires par les enseignants en fin de maternelle ou en début de 1ère année primaire prédisent mieux les performances académiques ultérieures ou le parcours scolaire des élèves que les évaluations cognitives et langagières standardisées. Ainsi, Quay et Steele (1998) montrent qu'une évaluation par les enseignants des performances académiques globales des élèves est prédictive de leurs performances ultérieures évaluées par une batterie

d'épreuves standardisées. Dans cette perspective, nous avons essayé d'étudier la valeur prédictive d'un questionnaire d'évaluation des comportements et compétences scolaires renseignés par les enseignants en classe préparatoire sur le niveau ultérieur en lecture. Nous cherchons ainsi à vérifier les résultats des études menées par Guimard & Florin (Florin, Guimard & Nocus, 2002 ; Guimard & Florin, 2007 ; Guimard, Cosnefroy & Florin, 2007 ; Guimard & Florin, 2001 ; Florin, 1991) en nous appuyant sur des données longitudinales obtenues auprès d'un échantillon d'élèves tunisiens.

L'examen de la littérature atteste que les évaluations cognitives et langagières standardisées ont une bonne valeur prédictive globale des performances scolaires ultérieures. En outre, ce type d'évaluation est assez fréquemment utilisé dans une perspective de prévention des difficultés d'apprentissage afin de repérer les enfants « à risque » et de mettre en œuvre les aides adaptées à leurs difficultés. De ce point de vue, nous étudierons la façon dont on peut passer d'une prédiction globale à une prédiction individuelle. De nombreuses recherches confirment la bonne valeur prédictive individuelle de certaines épreuves standardisées (Piquard-Kippfer, 2003 ; Peterson & Elbro, 1999) et des évaluations réalisées par les enseignants (Guimard & al., 2007). Il semble toutefois que les prédictions réalisées par les enseignants, tout comme celles issues des évaluations langagières et cognitives standardisées, sont nettement moins fiables lorsqu'elles concernent les futurs élèves en difficulté (Guimard & Florin, 2007 ; Piquard-Kipffer, 2003 ; Peterson & Elbro, 1999). Une des originalités de cette recherche est de s'interroger sur l'utilité de ces deux types d'évaluation pour la prévention des difficultés dans l'apprentissage de la lecture et de tester leur valeur prédictive individuelle. Il s'agit, d'autre part, d'étudier la capacité des évaluations réalisées par les enseignants et des évaluations cognitives et langagières à rendre compte des différences individuelles en lecture deux ans plus tard.

En langue française, de nouveaux outils d'évaluation sont disponibles pour déterminer le niveau des enfants dans différentes composantes cognitives ou langagières dans une perspective de diagnostic des troubles des apprentissages ou de prédiction/prévention des difficultés d'adaptation des enfants en contexte scolaire (Guimard & Florin, 2008). Parmi les batteries récentes, on peut citer le BSEDS[35], la BREV[36] et le Timé2[37]. En Tunisie, nous ne

[35] Bilan de Santé Evaluation du Développement pour la Scolarité 5 à 6 ans de Zorman et Jacquier-Roux (2001).
[36] Batterie rapide d'évaluation des fonctions cognitives de Billard et al. (2002).
[37] Test d'Identification de Mots Ecrits pour enfants de 6 à 8 ans de Ecalle (2002).

disposons de tels outils mais un projet de recherche en cours[38] a pour objectif la création d'outils de dépistage précoce des enfants présentant des facteurs de risque d'échec en lecture et éventuellement l'élaboration d'un diagnostic en vue de définir les perspectives de prise en charge. Ces épreuves, qui sont plus conformes au contexte et aux spécificités de la population tunisienne, sont en cours de validation. Le repérage des enfants risquant de se trouver en difficulté et l'identification de leurs difficultés spécifiques en fin de CP constituent les deux étapes nécessaires à une prise en charge précoce et efficace des difficultés dans l'apprentissage de la lecture. Il sera ainsi possible de prévenir les difficultés et de mettre en place des aides adaptées au développement des enfants.

2. Objectifs et hypothèses de la recherche

Le premier objectif de cette recherche consiste à déterminer les meilleurs prédicteurs de l'apprentissage de la lecture en langue arabe. De façon générale, les données de la littérature internationale montrent que les habiletés dans le domaine du langage oral et écrit sont significativement corrélées aux performances en lecture et semblent également prédire ces performances (*cf.* partie théorique, chapitre 4). L'évaluation des compétences langagières et cognitives d'un large effectif d'élèves tunisiens de CP[39] et le suivi de cette cohorte sur trois années scolaires (de la classe préparatoire à la 2ème année primaire) devraient permettre de mesurer le rôle et la part respective de ces compétences précoces dans l'explication de la réussite ou de l'échec dans l'apprentissage de la lecture.

La plupart des études menées sur les langues sémitiques ont permis d'établir certaines similitudes entre l'apprentissage de la lecture en langue arabe et en langues à écriture latines mais ont également mis en évidence des divergences liées aux caractéristiques de la langue. Nous essayerons d'apporter une lecture des résultats en rapport avec les spécificités de la langue arabe et du contexte tunisien.

Cet objectif conduit à formuler l'hypothèse suivante :

Hypothèse 1 : Les compétences langagières constitueraient de meilleurs prédicteurs de la réussite ou de l'échec dans l'apprentissage de la lecture en arabe que ne le sont les compétences cognitives.

[38] Programme de Recherche Fédéré (2007-2009) « Prise en charge des troubles des apprentissages et diagnostic différentiel avec le handicap mental ».
[39] L'équivalent de la GSM en France.

Nous formulons l'hypothèse que les compétences phonologiques et la connaissance des lettres seraient des prédicteurs puissants des performances ultérieures en identification de mots écrits en arabe. Concernant les compétences phonologiques, nous nous attendons à observer un effet de l'unité phonologique traitée dans la prédiction de l'apprentissage de la lecture. Ainsi, nous supposons que les compétences syllabiques et phonémiques expliqueraient une part importante de la variance en lecture et permettrait de prédire le niveau ultérieur en lecture. Toutefois, la sensibilité à la rime ne serait pas aussi importante pour l'identification de mots et ne semblerait pas prédire de façon significative le niveau ultérieur en lecture (identification de mots et compréhension). En effet, les études menées sur la langue arabe démontrent la non pertinence de la rime dans l'apprentissage de la lecture compte tenu de la situation de diglossie (Ammar, 2003 ; Boukadida, 2008 ; Besse, Demont & Gombert, 2007).

L'analyse des contraintes liées à la morphologie du mot et au système d'écriture de la langue arabe nous amène à nous interroger également sur le rôle de la conscience morphologique dans l'apprentissage de la lecture. Nous supposons que ces connaissances jouent un rôle important dans la prédiction de l'apprentissage de la lecture en arabe. Nous nous attacherons dans ce travail d'une part, à montrer l'existence d'un lien très fort entre les connaissances morphologiques précoces et le niveau ultérieur en lecture et d'autre part, à examiner le pouvoir prédictif de cette variable sur les performances en identification de mots mesurées à la fin de la première année et de la deuxième année de l'enseignement primaire. Nous nous attendons à ce que les connaissances morphologiques dérivationnelles expliquent une part significative de la variance en identification de mots et que les connaissances morphologiques flexionnelles contribuent de façon importante à la réussite en lecture-compréhension.

Les études récentes confirment également l'importance de la dénomination rapide ou RAN dans l'apprentissage de la lecture (Castel & al., 2008 ; Kirby & al., 2004). Si ce qui est mesuré par le RAN est bien l'accès rapide et automatique aux représentations phonologiques, nous supposons que les enfants tunisiens auront des difficultés à accéder de façon rapide et précise aux représentations phonologiques en arabe standard compte tenu de la situation de diglossie. On devrait alors observer de faibles prédictions des capacités de dénomination rapide sur le niveau de lecture en langue arabe.

Nous formulons également l'hypothèse que les habiletés cognitives non verbales ne seraient que faiblement reliées au futur niveau de lecture. La valeur prédictive de ces habiletés serait donc beaucoup moins puissante que celle des habiletés langagières précoces. Parmi ces

compétences, la mémoire à court terme apporterait la contribution la plus importante.
Nous pouvons également nous demander à quel aspect de la lecture tel ou tel prédicteur se trouve fortement associé. S'il a été suggéré que la conscience phonologique et morphologique (en particulier morpho-dérivationnelle) et la connaissance des lettres constituent les meilleurs prédicteurs de l'identification de mots, il serait raisonnable de supposer que les compétences morpho-syntaxiques, la compréhension de consignes et la compréhension orale constitueraient de bons prédicteurs de la lecture-compréhension.

Le second objectif de cette recherche consiste à proposer un modèle théorique des prédicteurs de l'apprentissage de la lecture en langue arabe et d'évaluer la qualité de l'ajustement de ce modèle aux données. Un modèle d'équations structurelles (MES) sera développé afin d'apprécier l'ensemble des relations prédictives entre les différentes variables de l'étude. La revue de la littérature a révélé en effet que les compétences langagières constituent des prédicteurs très puissants des capacités de lecture ultérieures (pour une synthèse voir Kirby, Desrochers, Roth & Lai, 2008). Les études qui ont essayé de modéliser la relation prédictive entre les compétences langagières précoces et l'apprentissage de la lecture signalent que certaines habiletés préscolaires ont davantage de poids que d'autres dans le développement de l'apprentissage de la lecture (Muter & al., 2004 ; Storch & Whitehurst, 2002). Les habiletés de sensibilité phonologique et les habiletés langagières sont reconnues comme des prédicteurs puissants de cet apprentissage. D'après le cadre théorique présenté, les connaissances phonémiques et morphologiques devaient avoir une forte incidence sur l'apprentissage de la lecture. Les premières permettent en effet d'identifier et de manipuler des éléments du langage oral (phonèmes) alors que les secondes permettent d'avoir accès aux plus petites unités sublexicales ayant un sens (morphèmes).

Dans le but de modéliser les relations prédictives entre les compétences langagières précoces et les performances ultérieures en lecture, nous proposons un modèle composé de quatre construits ou variables latentes : connaissances phonologiques, connaissances morphologiques, connaissances phonémiques et performances en lecture. Ainsi, nous proposons de tester l'hypothèse suivante :

Hypothèse 2 : Les connaissances phonologiques et morphologiques en classe préparatoire influenceraient positivement les performances en lecture en langue arabe en 2ème année primaire.

Nous postulons plus particulièrement que les connaissances phonologiques en classe préparatoire (CP) auraient un effet direct significatif sur les performances en lecture en 2ème année et un effet indirect par l'intermédiaire des connaissances phonémiques en 1ère année. En effet, de nombreuses études ont montré que la sensibilité aux unités phonologiques (syllabe, rime) permet l'émergence de la conscience phonémique qui a une incidence directe sur la lecture, tout comme la sensibilité aux unités syllabiques. Nous pouvons ainsi concevoir une continuité entre la sensibilité aux unités syllabiques et les capacités métaphonémiques. La sensibilité aux syllabes (implicite) serait un précurseur de la conscience phonémique (explicite) qui se développerait essentiellement sous l'effet de l'appretissage de la lecture. Les connaissances morphologiques en CP seraient associées directement et significativement au niveau de lecture en 2ème année.

Le modèle postulé des prédicteurs de l'apprentissage de la lecture en langue arabe serait en adéquation avec les données recueillies. Nous nous attendons à ce que la qualité d'ajustement du modèle théorique aux données soit satisfaisante selon les normes établies par Roussel, Durrieu, Compoy & El Akremi (2002).

Le troisième objectif de ce travail consiste à évaluer la contribution des évaluations réalisées par les enseignants dans la prédiction de la réussite dans l'apprentissage de la lecture en arabe. Les enseignants disposent en effet des compétences nécessaires pour observer avec précision le fonctionnement de leurs élèves à l'école et sont aussi en mesure de repérer certaines difficultés (Florin, 1991 ; Florin & al., 1999 ; Guimard & Florin, 2001). Leurs évaluations ou leurs pronostics s'avèrent prédictifs des difficultés scolaires des élèves dans la suite de la scolarité (Guimard & Florin, 2007 ; Guimard & al., 2007).

Suite à la revue de la littérature relative au rôle joué par les enseignants de maternelle dans le repérage des difficultés des élèves (*cf.* première partie, chapitre 4), nous pouvons envisager que les évaluations des compétences et comportements scolaires des élèves constituent des indicateurs particulièrement pertinents pour prédire leur niveau ultérieur en lecture. Dans ce sens, nous proposons la vérification de l'hypothèse suivante :

Hypothèse 3 : Les évaluations des comportements et des compétences scolaires réalisées par les enseignants en classe préparatoire seraient prédictives des performances ultérieures dans l'apprentissage de la lecture en langue arabe.

Nous supposons que les évaluations des comportements et des compétences scolaires réalisées par les enseignants en classe préparatoire permettraient de prédire la réussite ultérieure des élèves en lecture et prévenir leurs difficultés.

Dans le but de vérifier la fiabilité et la validité du questionnaire renseigné par les enseignants, nous avons adopté la modélisation par les équations structurelles. Nous proposons un modèle de mesure composé de trois facettes. Nous retenons les trois dimensions des évaluations des enseignants dégagées dans la littérature (Florin, Guimard & Nocus, 2002 ; Guimard & Florin, 2007 ; Guimard & Florin, 2001) : compétences langagières, comportements en classe et compétences scolaires ou transversales. Nous supposons que ce modèle de mesure, testé à partir de l'analyse factorielle confirmatoire, répondra aux qualités requises de fiablité et de validité interne.

Nous proposons également une première modélisation des relations prédictives entre ces différentes dimensions des évaluations des enseignants et les performances ultérieures en lecture. Une modélisation en équations structurelles permettra d'apprécier la nature des relations existant entre ces différentes variables. Nous supposons que le modèle théorique testé sera en adéquation avec les données empiriques. Nous nous attendons à ce que les évaluations des enseignants sur les aspects comportementaux soient plus liées aux performances en lecture que leurs évaluations sur les compétences langagières et transversales. En effet, de nombreux travaux ont mis en évidence que certains comportements des élèves en classe évalués par les enseignants au moyen de questionnaire (capacité à s'organiser dans les tâches scolaires, niveau de confiance en soi, etc.) constituent des indicateurs pertinents pour différencier les élèves et prédire leur évolution (pour une revue voir Guimard & Florin, 2007).

Le quatrième objectif consiste à mesurer la valeur prédictive individuelle des différentes mesures langagières et cognitives et des évaluations des enseignants. En d'autres termes, nous allons essayer de déterminer la proportion des normolecteurs et des enfants en difficulté de lecture correctement détectés par les mesures réalisées en classe préparatoire.

Plusieurs recherches ont mesuré la valeur prédictive individuelle des épreuves langagières et cognitives standardisées (Peterson & Elbro, 1999 ; Piquard-Kipffer, 2003) et de questionnaires d'évaluation renseignés par les enseignants (Guimard & Florin, 2007) afin de les utiliser dans une perspective de prévention. De façon générale, les résultats de ces

recherches montrent que leur valeur prédictive est moins pertinente lorsqu'il s'agit de prédire les difficultés ultérieures en lecture. Ainsi, les deux types d'évaluation parviennent à mieux prédire la réussite ultérieure en lecture que les difficultés. Ceci nous permet d'émettre l'hypothèse suivante :

Hypothèse 4 : les évaluations réalisées en classe préparatoire prédiraient mieux la réussite en lecture que les difficultés.

Nous nous attendons à ce que les évaluations réalisées par les enseignants et les compétences langagières et cognitives en classe préparatoire prédisent la réussite individuelle des élèves en lecture une année et deux années plus tard. Nous supposons en effet que ces évaluations auraient une bonne spécificité et une bonne valeur prédictive négative. En revanche, la qualité des prédictions concernant les futurs élèves en difficulté sera moins bonne. On s'attend à une sensibilité et une valeur prédictive positive moyenne ou faible.

Nous supposons également que les évaluations réalisées en classe préparatoire permetteraient de prédire les contrastes entre bons lecteurs, lecteurs moyens et faibles lecteurs que l'on peut observer au cours de la 2ème année primaire. Nous nous attacherons à vérifier que les différences individuelles observées à l'école primaire dans les domaines fortement liés à l'apprentissage de la lecture correspondraient en CP à des profils spécifiques.

3. Traitement statistique des données

Afin de vérifier nos hypothèses de recherche, plusieurs types d'analyses statistiques ont été menés.

Au cours du second chapitre de cette partie, nous avons analysé, dans un premier temps, les performances obtenues aux différentes mesures à T1, T2 et T3 (logiciel SPSS 16.0). Pour certaines épreuves, nous avons comparé la moyenne à la valeur du hasard à l'aide du test t pour échantillon unique. Dans un deuxième temps, des analyses de corrélation de Bravais Pearson et des analyses factorielles en composantes principales (ACP) avec rotation oblique ont été réalisées afin d'identifier les relations entre les différentes mesures à chaque niveau scolaire. Ces deux types d'analyse ont été menés séparément sur chaque niveau scolaire. Afin de vérifier la qualité de cohérence interne de chaque composante, nous avons eu recours au coefficient « Alpha de Cronbach » (α de Cronbach). Dans un troisième temps, dans le but d'évaluer la contribution et la valeur prédictive des différentes compétences langagières et

cognitives aux capacités ultérieures en lecture, des analyses de corrélation simple et ensuite des analyses de régression pas à pas (stepwise regression) ont été conduites.

Dans le troisième chapitre, nous avons eu recours à la modélisation par les équations structurelles afin d'établir un modèle des prédicteurs de l'apprentissage de la lecture en langue arabe à l'aide du logiciel AMOS 16.0. Cette technique statistique permet d'étudier les relations entre un ensemble de variables dont certaines sont considérées comme indépendantes et d'autres comme dépendantes. Ce type d'analyse permet de distinguer le type d'effets (positifs ou négatifs) des variables entre elles, d'estimer le poids des variables indépendantes sur les variables dépendantes et de déterminer par quelle voie (directe ou indirecte) chaque variable influence la suivante. En effet, le logiciel permet de tester l'ajustement du modèle théorique aux données. L'estimation de la qualité d'ajustement du modèle théorique aux données empirique est réalisée par la méthode du maximum de vraisemblance (Maximum Likelihood).

Dans le quatrième chapitre, nous avons adopté la même démarche que dans le second chapitre afin d'étudier la valeur prédictive des évaluations des comportements et compétences scolaires par les enseignants. Dans le but d'étudier plus finement la structure du questionnaire, une analyse de corrélation inter-items et une analyse de la fiabilité par l'alpha de Cronbach ont été réalisées. Le test de la validité du questionnaire est effectué par une analyse factorielle confirmatoire (AFC) réalisée avec le logiciel AMOS. Par ailleurs, la réalisation d'analyses de corrélation et de régression pas à pas a permis de déterminer d'une part, la nature des relations entre les différents items du questionnaire et les épreuves de lecture et d'autre part, la contribution de ces items à l'efficience ultérieure en lecture. Nous avons utilisé la méthode d'équations structurelles pour étudier les liens prédictifs entre les différentes dimensions du questionnaire renseigné par les enseignants et les performances en lecture. L'estimation du modèle est effectuée par la méthode du maximum de vraisemblance.

Enfin, nous nous sommes attachés dans le dernier chapitre de cette recherche à mesurer la valeur prédictive individuelle des différentes mesures langagières et cognitives et des évaluations des enseignants sur le niveau de lecture. Cette technique est généralement utilisée dans le domaine médical dans un objectif de dépistage ou de diagnostic des maladies. Les notions de sensibilité, de spécificité sont des paramètres statistiques permettant d'apprécier quantitativement la valeur du test. Les calculs de la sensibilité, de la spécificité, de la valeur prédictive positive et négative s'inspire des tests médico-statistiques permettant d'identifier les personnes malades et non malades. Dans la deuxième partie de ce chapitre, nous avons

effectué une analyse différentielle rétrospective. Les enfants ont été regroupés en trois groupes en fonction de leurs scores en lecture : bons lecteurs (scores supérieurs au centile 75), faibles lecteurs (scores inférieurs au centile 25) et lecteurs moyens (scores entre centiles 25 et 75). Pour chaque épreuve administrée en classe préparatoire, des analyses de variance (ANOVA) sur les performances obtenues par les trois groupes de lecteur ont été effectuées.

Chapitre 2 : Compétences évaluées en classe préparatoire et prédiction du niveau ultérieur en lecture

Un grand nombre d'études dans le domaine de l'apprentissage de la lecture tentent de définir les facteurs qui favorisent son acquisition. En effet, Il est aujourd'hui largement reconnu que la conscience phonologique joue un rôle déterminant dans l'apprentissage de la lecture et apparaît comme un bon prédicteur de la réussite ultérieure en lecture (Demont & Gombert, 2007 ; Ecalle & Magnan, 2007a ; 2007b ; Gombert, 1992 ; 2003 ; Gombert & al, 2004). Cependant, d'autres travaux ont mis en évidence le rôle déterminant d'autres compétences comme la connaissance des lettres et les capacités de dénomination rapide (Castel & al., 2008 ; Foulin, 2007 ; Lecoq, 1991 ; Schatschneider & al., 2004).

Sur le modèle de ce qui est établi à propos des liens existant entre l'apprentissage de la lecture et les compétences phonologique, des recherches de plus en plus nombreuses tentent de montrer que les connaissances morphologiques de l'enfant jouent un rôle important dans l'élaboration des procédures d'identification de mots, et ceci dès le début de l'apprentissage de la lecture (Boukadida, 2008 ; Colé & al., 2003 ; Marec-Breton, 2009 ; 2003 ; Marec-Breton & al., 2005). Les études menées sur la langue arabe confirment l'importance des connaissances morphologiques lors de la lecture de mots en arabe et montrent qu'elles expliquent une part importante de la variance en lecture (Besse & al., 2007 ; Belajouza, 2003 ; Boukadida, 2008).

Dans la présente recherche, nous essayerons de confirmer d'une part certains résultats déjà obtenus sur les langues à écriture latine et d'autre part, d'apprécier ce qui dépend des spécificités de la langue arabe et du contexte d'apprentissage.

1. Méthode

1.1. Sujets

L'échantillon initial était composé de 230 élèves tunisiens dont l'âge moyen était de 5 ans 7 mois lors des premières passations à la fin de la classe préparatoire (Mai-Juin 2005). Les enfants ont été recrutés dans 13 classes préparatoires appartenant à 10 écoles primaires

publiques de la région de Tunis. Les enfants, sélectionnés sur la base de leurs dossiers scolaires[40], ont un niveau socio-économique moyen.

Tableau 5 : Caractéristiques de la population

Niveau	Effectif	Age moyen (en mois)	Ecart-type
Classe préparatoire (CP)[41]	230	67.70	3,78
1ère année[42]	184	80.57	3,52
2ème année[43]	171	93.67	3,09

Les enfants ont été évalués à trois reprises où différentes épreuves leur ont été administrées : une première fois en Mai-Juin 2005 (T1), une deuxième fois en Mai 2006 (T2) et une dernière fois en Mai 2007 (T3). Entre la fin de la classe préparatoire et la fin de 1ère année, 56 enfants ont quitté l'étude pour diverses raisons (déménagement, changement d'école, passage dans l'enseignement privé, etc.). En fin de 2ème année primaire, 171 élèves, 80 garçons et 91 filles d'âge moyen 7 ans 8 mois, fréquentaient toujours l'établissement scolaire dans lequel ils avaient été testés en 1ère année. Les participants sont tous de langue maternelle arabe et aucun n'avait de problèmes particuliers (déficit sensoriel, trouble du langage). Les analyses ont porté sur l'échantillon final (N=171) pour lequel nous avons toutes les données.

1.2. Matériel

1.2.1. Mesures proposées en classe préparatoire (T1)

Au cours de la première évaluation, différentes épreuves langagières et cognitives ont été administrées à des élèves en classe préparatoire lors de passations individuelles ou collectives. En ce qui concerne la familiarité du matériel proposé, nous avons procédé à l'examen des livres de lecture des enfants pour chaque niveau scolaire et nous nous sommes référés à l'avis des enseignants.

1.2.1.1. Evaluation du langage oral

a. Connaissances phonologiques

La majorité des chercheurs s'accordent sur l'importance du développement de la conscience phonologique dans l'apprentissage de la lecture (Goswami & Bryant, 1990 ; Gombert, 1990 ;

[40] Le dossier scolaire de chaque enfant mentionne le niveau socio-économique de la famille.
[41] L'équivalent de la GSM dans le système éducatif français.
[42] L'équivalent du CP dans le système éducatif français.
[43] L'équivalent du CE1 dans le système éducatif français.

Gombert & al, 2004 ; 1992 ; Muter & al., 1998). Plusieurs études ont montré que les performances des enfants dans plusieurs épreuves mesurant leurs capacités métaphonologiques étaient prédictives du niveau de lecture ultérieur (Muter & al., 1998 ; Muter & al., 2004 ; Sprenger-Charolles & al., 2000 ; Stanovich & al., 1984 ; Storch & Whitehurst, 2002).

Dans le but d'évaluer les connaissances phonologiques des enfants de classe préparatoire, nous avons utilisé différentes tâches : détection de l'intrus phonologique (syllabe, rime et phonème) en passation collective, comptage et segmentation syllabique et suppression syllabique en passation individuelle. Nous avons essayé ainsi de varier le type de la tâche (détection de l'intrus, comptage, segmentation et suppression), les types d'unités à prendre en considération (rime, syllabe, phonème) et la place des unités à traiter dans les mots (début, milieu, fin). La formulation des consignes de chaque épreuve est adaptée au niveau de compréhension des enfants. Les consignes ont été énoncées en arabe dialectal et nous avons évité les mots techniques comme syllabe ou phonème (*cf.* annexe I). Dans chaque tâche, la phase d'évaluation est précédée d'une phase d'entraînement avec feed-back correctif.

Détection de l'intrus de rime

Cette épreuve vise à évaluer la sensibilité des jeunes enfants à la structure sonore des mots. Le matériel est constitué de 6 listes de trois mots monosyllabiques comportant un intrus de rime (*cf.* annexe I). L'enfant doit choisir parmi ces trois mots, celui qui ne rime pas avec les deux autres (i.e., فار - ساق - كار). Un point est attribué à chaque réponse correcte.

Détection de l'intrus syllabique

Pour la détection de la syllabe initiale, nous avons suivi la même procédure que dans le cas précédent. Le matériel est constitué de 6 listes de trois mots comportant un intrus de syllabe initiale (*cf.* annexe I). La tâche consistait à dire lequel des trois mots commençait différemment des autre (i.e., ورقة - وردة - معزة). Un point est attribué à chaque réponse correcte.

Comptage et segmentation syllabique

Cette épreuve qui comporte 12 items mesure la capacité des élèves à prendre conscience de la structure syllabique des mots. L'exercice consiste à demander à l'enfant de décomposer, à l'oral, des mots comportant une, deux ou trois syllabes. Dans un premier temps, l'enfant doit préciser le nombre de segments dans chaque mot avec ses doigts ensuite, il doit les nommer

(i.e., nous demandons à l'enfant de compter le nombre de syllabes dans le mot صابون [ṣa:bu:n][44] ensuite de les nommer). Deux points sont attribués à chaque réponse correcte. Si l'enfant échoue à l'une des deux tâches, on lui attribue 1 point.

Suppression syllabique

L'exercice consiste à demander à l'enfant de dire ce qui reste d'un mot lorsqu'on en supprime une syllabe. Les mots sont composés de deux ou trois syllabes (*cf.* annexe I). La position de la syllabe à supprimer peut être au début, au milieu ou à la fin du mot (i.e., nous demandons à l'enfant de dire ce qui reste du mot حَلَزون [ħalazu:n][45] si on retire la syllabe [ħa]). Un point est attribué à chaque réponse correcte.

Détection de l'intrus phonémique

La majorité des études s'accordent sur l'importance du développement de la conscience phonologique dans l'apprentissage de la lecture et essentiellement la conscience phonémique qui serait le facteur le plus important dans cet apprentissage (Muter & al., 2004 ; Muter & al., 1998). Cette épreuve, administrée en passation collective, est destinée à évaluer la sensibilité phonémique des enfants. Elle comprend un total de 12 séries de trois mots chacune (6 items comportant un intrus de phonème initial et 6 items comportant un intrus de phonème final). Chaque item est composé de trois mots représentés par des images. Dans la tâche de détection du phonème final, les enfants devaient entourer l'image du mot qui ne finit pas comme les deux autres (i.e., ساق - بوق - فيل). Dans la tâche de détection du phonème initial, ils devaient entourer l'image du mot qui ne commence pas de la même façon que les deux autres (i.e., جُبن - بَرق - جِسر). Un point est attribué à chaque réponse correcte.

b. Connaissances morphologiques

Sur le modèle de ce qui est établi à propos des liens existant entre l'apprentissage de la lecture et les compétences métaphonologiques, des études de plus en plus nombreuses tentent de montrer qu'une sensibilité morphologique est nécessaire à l'apprentissage de la lecture (Carlisle, 1995 ; Marec-Breton, 2003 ; 2009). Le niveau de connaissances morphologiques des enfants a été évalué par trois tâches administrées en passation collective : deux tâches de morphologie flexionnelle et une tâche de morphologie dérivationnelle (*cf.* annexe I).

[44] Savon en français
[45] Escargot en français

Jugement de l'ordre des mots

Cette tâche vise à évaluer la sensibilité des sujets à l'ordre des mots dans une phrase. Le matériel est constitué de 10 énoncés : 5 énoncés sont grammaticalement corrects (i.e., الوليّد روّح من المكتب) et 5 énoncés sont incorrects (i.e., العصفور واقف الشجرة فوق). Les fautes se situent au niveau de l'ordre des mots dans la phrase (*cf.* annexe I). La tâche de l'enfant consiste à entourer une image s'il trouve que la phrase énoncée est correcte et faire une croix s'il trouve qu'elle est incorrecte. Un point est attribué à chaque bonne réponse.

Jugement morpho-syntaxique

Cette tâche vise à évaluer la capacité de raisonnement des enfants sur les aspects morpholo-syntaxiques. Le matériel est constitué de 10 items dont 5 sont grammaticalement corrects (i.e., سامي شرب كاس حليب) et 5 grammaticalement incorrects (i.e., أحمد كلات تفّاحة). Les erreurs se situent au niveau de la morphologie des verbes (mauvais accord du verbe avec le sujet). La tâche de l'enfant consiste à entourer une image s'il trouve que la phrase énoncée est correcte et faire une croix s'il trouve qu'elle est incorrecte. Un point est attribué à chaque bonne réponse.

Jugement des liens morphologiques

La présente épreuve a été conçue dans le cadre d'un projet de recherche national sur la prévention des troubles des apprentissages scolaires (Projet de Recherche Fédéré)[46]. L'épreuve utilisée pour mesurer le niveau de sensibilité morphologique des enfants est composée de 16 séries de deux mots. Les sujets doivent se prononcer sur l'existence ou non d'un lien morphologique entre les deux mots présentés oralement par l'expérimentateur (les enfants doivent dire s'ils sont de la même famille ou non). Le matériel est composé de 16 paires de mots : 8 appartiennent à la même famille morphologique (i.e., حيوان - حيوانات), 4 présentent un lien sémantique (i.e., مربَع - مثلّث) et 4 présentent un lien phonologique (i.e., كرسي - كرَاسة). On s'assure que l'enfant a compris le concept de « famille morphologique » en lui fournissant des exemples des trois types d'items. Un point est attribué à chaque réponse correcte.

c. Dénomination rapide

Lire suppose, non seulement que l'information relative à la langue soit stockée en mémoire, mais aussi qu'elle soit facilement et rapidement accessible (Castel & al., 2008). Cette facilité

[46] Cette épreuve a été conçue au sein de l'équipe consacrée au langage oral et écrit dont je fais partie.

d'accès peut être évaluée à partir de la durée nécessaire à la dénomination. L'épreuve utilisée a été réalisée par une équipe de chercheurs tunisiens dans le cadre d'un projet de recherche (modèles d'acquisition de la lecture et évaluation de la compétence lexique : l'apport de l'arabe aux études inter-langues) (travaux non publiés). Cette épreuve teste la rapidité et la précision du sujet à identifier les images et à récupérer le mot correspondant avec une contrainte de temps. Il s'agit donc d'évaluer l'accès au lexique mental. Dans cette épreuve, on demande à l'enfant de dénommer, le plus rapidement et le plus correctement possible, une série de 32 items imagés en arabe dialectal. Les images représentent des objets et des animaux familiers à l'enfant et ne posent donc aucun problème d'identification. Le calcul du score tient compte de la durée et de la précision de la dénomination de toutes les images.

d. Compréhension orale

Outre le décodage, la lecture suppose un contrôle des traitements menant à la compréhension (Gombert & Colé, 2000). Dans le but de mesurer la compétence des sujets à dégager la signification d'énoncés oraux, nous avons utilisé une épreuve de 20 énoncés adaptée de l'épreuve E.20 de Khomsi (1990) dans le cadre d'un travail de recherche réalisé par une équipe de chercheur de l'université de Tunis et de l'université de Nantes (travaux non publiés). Cette épreuve, administrée en passation collective, cherche à identifier les stratégies de compréhension d'énoncés en contexte imagé. Elle consiste à énoncer oralement des phrases en arabe dialectal tunisien et l'élève doit choisir, parmi un ensemble de quatre images proposées, celle qui exprime le mieux le sens de la phrase énoncée. Un point est attribué à chaque réponse correcte.

e. Compréhension de consignes

Un des aspects de la compréhension du langage oral est la compréhension de consignes. Cette compétence n'est pas habituellement prise en compte dans les recherches sur l'apprentissage de la lecture mais nous avons jugé qu'elle constitue une des conditions de la réussite des enfants dans les tâches scolaires. En effet, la majorité de ces tâches est introduite par des consignes plus ou mois complexes. Nous avons utilisé une épreuve proposée par le ministère de l'éducation nationale français (2001)[47] que nous avons adapté au contexte tunisien. Le support utilisé se présente sous la forme d'une image représentant une scène. L'objectif est d'amener l'enfant à réaliser une ou plusieurs actions précises (colorier, dessiner, entourer).

[47] Il s'agit d'une première série d'outils d'évaluation et d'aide aux apprentissages à l'entrée de la la grande section de maternelle et à l'entrée du cours préparatoire dans différents domaines.

L'évaluation se déroule en passation collective. Chaque élève dispose d'une feuille sur laquelle figure l'image et des crayons de couleurs. Six consignes de difficulté croissante sont énoncées à l'enfant. Un point est attribué lorsque la consigne est correctement excécutée.

1.2.1.2. Evaluation du langage écrit

a. Connaissance des lettres

La connaissance de l'alphabet repose sur un savoir implicite relatif à la structure phonologique du nom des lettres. Ainsi, l'apprentissage des CGP est plus facile lorsque le nom de la lettre est acquis puisque l'établissement de la relation son-lettre peut s'appuyer sur le nom de la lettre contenant le phonème à mémoriser (Ecalle & Magnan, 2002 ; Biot-chevrier & al., 2008). Dans cette épreuve, l'enfant doit reconnaître des lettres présentées à l'oral sur une feuille parmi d'autres items écrits et les entourer. Le matériel est constitué de 12 listes d'items comportant chacune un item cible à reconnaître parmi des items distracteurs. Les items distracteurs sont : une lettre proche visuellement ou phonologiquement, une lettre sans lien apparent avec la lettre cible, un chiffre et un symbole graphique (*cf.* annexe I). Un point est attribué à chaque réponse correcte.

1.2.1.3. Evaluation des compétences cognitives

a. Intelligence non verbale

L'épreuve utilisée pour évaluer le niveau intellectuel non verbal des enfants est le Test d'Aptitude Non verbale de Naglieri (1998) (forme A). Le NNAT est un test d'intelligence non verbale basé sur le principe des matrices analogiques. Il permet une évaluation fiable des aptitudes au raisonnement non verbal et à la résolution de problèmes quels que soient le niveau socioculturel et les compétences verbales du sujet. Il est composé de 38 items de difficulté croissante présentés sous formes d'une série de figures principale auxquels il manque une partie. La tâche du sujet est d'identifier, parmi les cinq petites figures proposées en bas, celle qui complète le mieux la figure-stimulus. Le NNAT est administré en passation individuelle d'une durée moyenne de 15 minutes. Un score global est calculé correspondant à la somme des bonnes réponses.

b. Attention

De nombreuses recherches ont démontré que la lecture chez le lecteur débutant nécessite une mobilisation d'une attention soutenue pour exécuter les différentes opérations cognitives impliquées dans cette activité mentale (Alegria & Morais, 1996 ; Content, 1996 ; Valdois,

2003). L'épreuve utilisée pour l'évaluation des capacités attentionnelles des enfants est le test des deux barrages de Zazzo (1967). L'enfant doit identifier dans une suite de signes, ne différant que par un seul élément, les signes identiques au modèle proposé et les barrer le plus rapidement possible. Ce test comporte deux phases : une première pendant laquelle il s'agit de barrer un seul type de signes et une seconde où il s'agit de barrer deux types de signes. Le temps de passation a été fixé à 5 minutes pour tous les enfants. Le score à l'épreuve correspond au quotient des rendements R2/R1 (R1 : rendement au barrage d'un signe ; R2 : rendement au barrage de deux signes). Le calcul de ces indices a été effectué en se reportant aux tableaux du manuel pour l'examen psychologique de l'enfant (Zazzo, 1967).

c. Mémoire à court terme (MCT)

Plusieurs recherches (Demont, 2002 ; Lecoq & al., 1996) ont démontré que la capacité mnésique est un facteur important dans l'explication des performances en lecture. Dans l'activité de lecture, la mémoire immédiate joue un rôle très important dans le stockage momentané et le traitement des informations relatives aux mots et à la représentation propositionnelle de l'énoncé.

Pour évaluer la capacité mnésique à court terme, nous avons utilisé une tâche inspirée du test de mémoire à court terme crée au sein du laboratoire « Labécd » de l'université de Nantes (2001). L'épreuve est administrée en passation collective. Pour cela, chaque enfant dispose d'un petit carnet avec des dessins sur chaque page, une page blanche entre chaque dessin. L'enfant doit mémoriser les mots qu'il entend (deux séries croissantes de 2 à 5 mots) puis retrouver les images correspondantes à ces mots sur le carnet et les entourer. Un point est attribué à chaque série correctement mémorisée.

d. Concepts de base

Selon Boehm (1990), la connaissance des concepts de base est nécessaire aux enfants pour se conformer aux instructions de l'enseignant, comprendre le contenu du matériel éducatif et communiquer avec les autres. Il s'agit de mesurer le degré de maîtrise de quelques notions de temps et d'espace considérées comme fondamentales pour la réussite des premières années de scolarité. La tâche consiste à entourer un dessin correspondant à la consigne donnée par l'expérimentateur dans une série de trois ou quatre dessins. Les dessins proposés sont empruntés au Test des Concepts de Base de Boehm (1990) en version préscolaire. A chaque dessin est associée une tâche qui implique la connaissance d'une notion temporelle ou

spatiale. Au total, l'épreuve est composée de 13 items expérimentaux et d'un item d'entraînement. Chaque item est destiné à mesurer une notion particulière liée au temps ou à l'espace. Les notions évaluées sont : *droite* (item 2), *gauche* (item 4), *avant* (item 11), *après* (item 7), *commence* (item 6), *entre* (item 3), *fini* (item 10), *devant* (item 1), *au milieu* (item 8), *au dessous* (item 5), *le plus petit* (item 9), *le plus haut* (item 13), *le plus long* (item 12). Un score global est calculé à partir des notes obtenues aux différents items.

e. Graphisme

L'épreuve utilisée pour évaluer les capacités graphiques est l'épreuve graphique d'organisation perceptive pour enfants de 4 à 6 ans de Santucci reprise par Zazzo dans son manuel pour l'examen psychologique de l'enfant (Zazzo, 1967). La tâche de l'enfant consiste à copier des formes géométriques de complexité variable présentes devant lui. L'épreuve est composée de neuf figures qui mettent en jeu différents rapports spatiaux. Un point est attribué à chaque réponse correcte selon un barème proposé par l'auteur.

1.2.2. Mesures proposées en 1ère année primaire (T2)

a. Identification de mots écrits

La plupart des recherches montrent très clairement que les difficultés des faibles lecteurs se situent au niveau des mécanismes de reconnaissance des mots. L'épreuve utilisée pour évaluer les performances en lecture est l'épreuve de l'identification des mots (ECIM-E) créée par Khomsi (1994) et adapté au contexte tunisien par une équipe de chercheurs tunisiens (travaux non publiés). Cette épreuve permet d'évaluer le niveau de compétence en identification des mots et de mettre en évidence les stratégies utilisées par le lecteur. Il s'agit d'une épreuve de lecture en contexte imagé dans laquelle il est demandé au sujet de juger de l'adéquation entre un mot et une image. Les enfants doivent accepter ou rejeter les items proposés selon que la dénomination est congruente ou non avec l'image ou que les mots sont correctement orthographiés ou non. L'épreuve de 60 items a été présentée en version vocalisée (*cf.* annexe III). Elle comporte 4 types d'items présentés de façon aléatoire :

- 20 items congruents avec les images et correctement orthographiés (OC).
- 10 items non congruents avec les images (PS pour Pseudo-synonymes).
- 15 items congruents avec les images mais ayant subis des perturbations orthographiques sur les consonnes (PLC).

- 15 items congruents avec les images mais ayant subis des perturbations orthographiques au niveau des signes vocaliques (PLV).

Les différents items ont été choisis en fonction de leur familiarité et de leur proximité avec leurs équivalents en arabe dialectal. La passation de l'épreuve est précédée d'une phase d'entraînement, au cours de laquelle, on présente à l'enfant quatre items différents. Un score global est calculé pour chaque enfant correspondant au nombre de réponses correctes.

1.2.3. Mesures proposés en 2ème année primaire (T3)

Pour évaluer le niveau de lecture en 2ème année, nous avons administrés aux enfants, en plus de l'épreuve d'identification de mots utilisés en 1ère année, une autre épreuve d'identification de mots écrits et une épreuve de compréhension.

a. Test d'Identification de Mots Ecrit de 6 à 8 ans (Timé2)

Pour évaluer le niveau de compétence en identification de mots en 2ème année, nous avons utilisé un test d'identification de mots inspiré du Timé2 (Ecalle, 2003) (*cf.* annexe III). Il se présente sous la forme de trois tâches distinctes de 12 items proposées en passation collective : reconnaissance de mots fournis oralement par l'expérimentateur (Timé2.1), reconnaissance de mots écrits, l'enfant devant choisir parmi cinq mots celui correspondant à une image (Timé2.2) et catégorisation sémantique consistant à choisir un mot sémantiquement ou pragmatiquement proche d'un mot inducteur (Timé2.3). Dans les trois tâches du Timé2, un mot cible est à rechercher dans une liste composée de 5 items tests : le mot correctement orthographié (شَمْسٌ [ʃamsun])[48], un pseudo-mot phonologiquement plausible (شَمْصٌ [ʃamsˁun]), un pseudo-mot visuellement proche (شَمَسٌ [ʃamasun]), un voisin orthographique (شَمْعٌ [ʃamʔˁun]) et un pseudo-mot comportant une séquence illégale de lettres (شْسَمٌ [ʃsamun]). Une note globale sur 36 points correspondant à la somme des notes composites (note sur 12) résumant les compétences des élèves en identification de mots a été calculée pour chaque enfant.

b. Epreuve de compréhension écrite

Nous avons utilisé l'épreuve de compréhension écrite adaptée de l'épreuve E.20 (Khomsi, 1990)[49]. Cette épreuve cherche à identifier les stratégies de compréhension d'énoncés en

[48] Soleil en français.
[49] Par une équipe du département de psychologie de la Faculté des sciences Humaines et sociales de Tunis et du laboratoire « Labécd » de l'Université de Nantes (travaux non publiés).

situation de contexte-image. L'enfant doit choisir l'image congruente avec un énoncé parmi un ensemble de quatre images. Les énoncés cherchaient à couvrir un ensemble de formes syntaxiques caractéristiques de la langue arabe (phrase nominales et phrases verbales, structures de l'accompli et de l'inaccompli…). L'épreuve comporte 16 items présentés sous forme vocalisée. Un point est attribué à chaque réponse correcte.

1.3. Procédure

En classe prépararoire (T1), différentes épreuves langagières et cognitives ont été présentées aux enfants lors de passations semi-collectives (6-7 élèves) et individuelles. Concernant l'évaluation du langage oral et écrit, nous avons utilisé certaines épreuves qui existent déjà (dénomination rapide, compréhension orale), alors que d'autres ont été construites en tenant compte à la fois de la spécificité de la langue arabe, du contexte sociolinguistique tunisien et des caractéristiques du système éducatif dans notre pays. Toutes les expériences se sont déroulées à la fin de l'année scolaire (entre mai et juin), c'est-à-dire au terme de 8 à 9 mois d'enseignement. Chaque passation est réaliseé dans une salle de l'école tenue à notre disposition.

En raison de la situation de diglossie, les consignes ont été énoncées en arabe dialectal. En classe préparatoire, tous les items des épreuves d'évaluation du langage oral étaient en arabe dialectal. Ainsi, nous avons essayé d'utiliser des épreuves adaptées et accessibles à des enfants arabophones qui n'ont pas encore appris à lire. Nous avons introduit l'arabe standard en 1ère et e 2ème année. Compte tenu du nombre important d'épreuves, trois séances se sont avérées nécessaires. Dans un premier temps, les enfants ont passé, de façon individuelle, l'épreuve d'intelligence non verbale, les tâches de comptage, segmentation et suppression syllabique et l'épreuve de dénomination rapide. Au cours d'une seconde séance, nous avons administré aux enfants, de façon collective, les épreuves de détection de l'intrus syllabique, de morpho-syntaxe, de graphisme, de MCT et de compréhension orale. Enfin, les enfants ont passé les épreuves de concepts de base, d'attention, de connaissance des lettres, de détection de l'intrus phonémique, de jugement de liens morphologiques et de compréhension de consignes. L'ordre de passation est le même pour tous les enfants.

Etant donné l'âge des enfants, une seule séance est programmée par jour pour chaque groupe pour éviter la fatigue et l'absence de collaboration des élèves. Chaque séance dure en moyenne 35 minutes. Afin que les enfants comprennent ce qui leur est demandé, chaque épreuve est précédée d'une phase d'entraînement avec feed-back correctif. Certaines épreuves

se sont avérées difficiles pour les enfants prélecteurs (effet plancher), donc nous avons décidé de les repasser en 1ère année. En fin de 1ère année primaire (T2), le niveau de lecture a été évalué à l'aide d'une épreuve collective d'identification de mots écrits en contexte imagé. Nous avons proposé également les épreuves pour lesquelles nous avons obtenu un effet plancher en CP : les épreuves de détection d'intrus phonémique (initial et fianl) et l'épreuve de jugement de liens morphologiques. En fin de 2ème année primaire (T3), les enfants ont passé deux nouvelles épreuves de lecture en passation collective : un test d'identification de mots écrits (Timé2) et une épreuve de compréhension écrite.

2. Résultats

2.1. Analyse des performances à T1, T2 et T3

Les performances aux différentes épreuves ont fait l'objet dans un premier temps d'analyses descriptives (N=171).

2.1.1. Performances aux épreuves langagières en classe préparatoire (T1)

Le tableau 6 présente les résultats obtenus aux différentes épreuves langagières à T1 (moyenne, écart-type, minima et maxima). Les résultats montrent qu'il existe une large dispersion dans les différentes épreuves utilisées. Certains enfants ont des performances faibles, d'autres montrent une performance moyenne et un grand nombre ont une performance intermédiaire.

Concernant les épreuves individuelles de manipulation des unités phonologiques, nous constatons que la tâche de comptage syllabique est réussie avec un score moyen de 66,6% (σ=3,8) et celle de segmentation syllabique avec un score moyen de 60% (σ=4,9). L'épreuve de suppression syllabique, impliquant un plus haut niveau de conscience phonologique, s'est avérée particulièrement difficile pour les enfants avec un taux de réussite de 37,8% (σ=3,9). En effet, l'analyse des fréquences montre que 18% des enfants n'ont donné aucune réponse correcte à cette épreuve. L'analyse de la variance met en évidence un effet du type de la tâche ($F(2,170)=27,8$; $p<.001$).

Pour les tâches de détection de l'intrus syllabique, nous constatons que les élèves de CP réussissent la tâche détection de rime à 57% (σ=1,46) et la tâche de détection de l'intrus syllabique à 65% (σ= 1,45). Les valeurs aux tests t font apparaître que les résultats aux deux

épreuves se distinguent significativement du hasard (respectivement t(1,170)= 27,15 ; p<.001 et t(1,170)= 33,10 ; p<.001). Toutefois, les scores aux épreuves phonémiques font appraitre des effets planchers se traduisant par des performances quasi nulles dans les tâches de détection de l'intrus phonémique (14%). Ces résultats mettent en évidence la difficulté que présente l'analyse phonémique par rapport à l'analyse syllabique.

Tableau 6 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus aux épreuves langagières à T1

Mesures	Moyenne (% moyen de réussite)	Ecart-type	Minimum	Maximum
Compt syllabique (/12)	8 (66,6%)	3,82	0	12
Segment syllabique (/12)	7,29 (60,7%)	4,93	0	12
Suppr syllabique (/12)	4,54 (37,8%)	3,92	0	11
Détect de l'intr syll (/6)	3,91 (65,1%)	1,46	1	6
Détect de l'intr rime (/6)	3,42 (57%)	1,45	0	6
Détect de l'intr phoné (/12)	1,7 (14%)	2,34	0	2
Morpho-syntaxe (/10)	5,26 (52,6%)	1,96	0	10
Jug de l'ordre (/10)	5,64 (56,4%)	2,12	2	10
Jug liens morph (/20)	6,75 (33,7%)	2,60	0	8
Conn des lettres (/12)	7,12 (59,3%)	3,39	0	12
Compréh orale (/16)	8,64 (54%)	2,20	3	14
Compréh de consignes (/6)	4,40 (73,3%)	1,25	1	6
Déno rapide	87 (40,4%)	36	40	215

Les résultats aux épreuves morpho-flexionnelles (connaissances morpho-syntaxiques et jugement de l'ordre) se révèlent plutôt moyens (respectivement 52,6% et de 56,4%), ce qui témoigne d'une certaine connaissance des règles morpho-syntaxiques permettant d'identifier la segmentation grammaticale des mots composant une phrase même si l'échantillon était composé d'enfants pré-lecteurs. Ces résultats se différencient positivement des réponses données au hasard (t(1,170)=3,96 ; p<.05 et t(1,170)=1,91 ; p<.05). Cependant, les élèves rencontrent plus de difficultés dans la tâche morpho-dérivationnelle avec un taux de réussite de 33% et un écart type de 2,60. Cette valeur ne diffère pas du niveau du hasard. Pour l'épreuve de connaissance des lettres, 59% des élèves répondent correctement. Là encore ce résultat montre qu'à ce niveau scolaire, les enfants pré-lecteurs sont capables d'identifier certaines lettres de l'alphabet. Pour l'épreuve de compréhension orale et de compréhension de

consignes, les enfants obtiennent respectivement 56,2% (σ=2,2) et 73,3% (σ=1,25) de réponses correctes. Enfin, les résultats obtenus à l'épreuve de dénomination rapide sont faibles (40,4% de réponses correctes).

2.1.2. Performances aux épreuves cognitives en classe préparatoire (T1)

Le tableau 7 indique les scores moyens de réussite aux différentes épreuves cognitives qui varient entre 33% et 81%. L'épreuve d'orientation spatio-temporelle obtient les meilleures performances avec un pourcentage moyen de réussite d'environ 81% (σ=1,7). Les élèves ont bien réussi l'épreuve de graphisme (68,77% ; σ= 1,8) et un peu moins bien la tâche de mémoire à court terme (51,62% ; σ=1,8). Nous remarquons la dispersion peu élevée des notes autour de la moyenne.

Tableau 7 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus aux épreuves cognitives à T1

Mesures	**Moyenne** **%moyen de réussite**	**Ecart-type**	**Minimum**	**Maximum**
Intelligence (/38)	12,87 (33,86%)	5	1	29
MCT (/8)	4,13 (51,62%)	1,8	0	8
Graphisme (/9)	6,19 (68,77%)	1,8	1	9
Concepts de base (/13)	10,53 (81%)	1,7	5	13
Attention	36,74 (40,37%)	16,1	1	91

Les épreuves d'attention et d'intelligence présentent des performances moins élevées avec des taux de réussite respectifs de 40,37% (σ= 16,1) et 33,86% (σ=5).

2.1.3. Performances aux épreuves langagières en 1ère année (T2)

Les performances obtenues par les enfants aux différentes épreuves langagières à T2 sont présentées dans le tableau 8. Rappelons que ces épreuves ont été administrées en classe préparatoire mais nous avons été confrontés à un effet « plancher ». En effet, les scores obtenus par les enfants de CP aux épreuves de détection de l'intrus phonémique et de jugement de liens morphologiques ne se distinguent pas du hasard, ce qui laisse penser que ces compétences se développent avec l'apprentissage de la lecture.

Vu l'importance de ces compétences dans l'acquisition de la lecture, nous avons jugé plus adéquat de les réadminister en 1ère année (T2) afin d'examiner leur pouvoir prédictif sur le

niveau de lecture en 2ème année (T3).

Tableau 8 : Performances moyennes de réussite, écart-types, minima, maxima obtenus aux épreuves langagières à T2

Mesures	Moyenne	Ecart-type	Minimum	Maximum
Détect du phonè initial (/6)	3,69	1,78	0	6
Détect du phonè final (/6)	3,19	1,74	0	6
Jug des liens morph (/20)	13	3,6	11	20

Les taux de réussite relevés dans les épreuves langagières administrées en 1ère année (T2) sont présentés dans la figure 7.

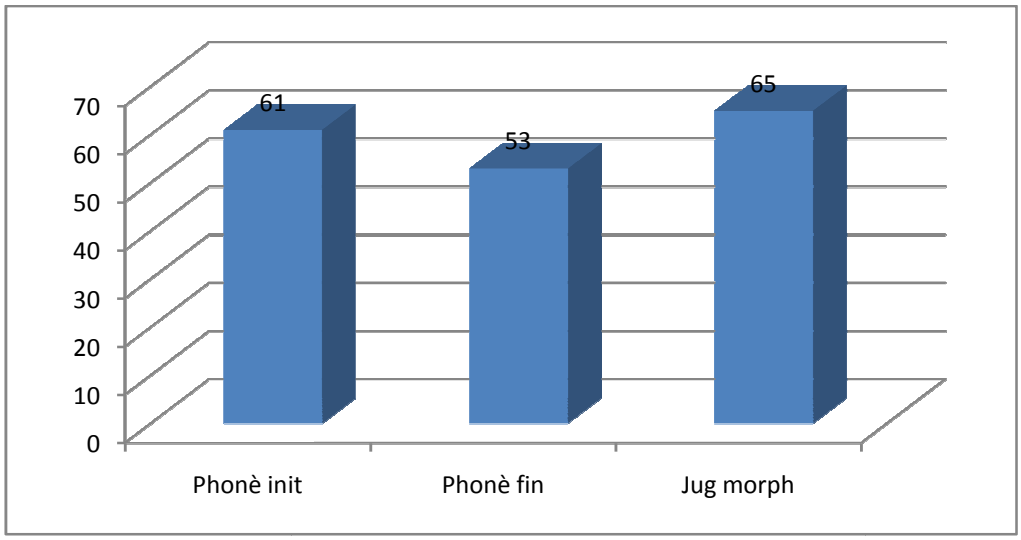


Figure 7 : Scores moyens (en %) obtenus aux différentes épreuves langagières à T1

Les résultats indiquent qu'à ce niveau, les enfants sont capables de manipuler les phonèmes, avec 61% de réussite dans la détection du phonème initial (σ=1,78) et 53% dans la détection du phonème final (σ=1,74). Les valeurs aux tests t font apparaître que ces résultats se distinguent significativement du hasard (respectivement t(1,170)= 8,87 et t(1,170)= 12,37 ; p<.05). Un niveau de réussite moyen est obtenu également à l'épreuve de jugement de liens morphologiques avec un taux de réussite d'environ 65%. Ce résultat se différencie significativement du niveau de hasard (t(1,170)= 19,77 ; p<.05). Nous constatons ainsi que les enfants, en 1ère année, commencent à intégrer la notion de famille morphologique.

L'enfant dispose donc au début de l'apprentissage de la lecture de connaissances phonémiques et morphologiques (dérivationnelles) non négligeables.

2.1.4. Performances aux épreuves de lecture à T2 et T3

2.1.4.1. Les performances en identification de mots à T2

Le tableau 9 récapitule les performances obtenues par les enfants à l'épreuve d'identification de mots à T2 (IME1).

Tableau 9 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus à l'épreuve de lecture IME1[50]

Mesures	**Moyenne (% moyen de réussite)**	**Ecart-type**	**Minimum**	**Maximum**
SG (sur 60)	42,14 (70,23%)	8,2	19	56
OC (sur 20)[51]	16,28 (81,4%)	5,16	10	20
PLC (sur 15)[52]	10,15 (67,66%)	3,07	0	15
PLV (sur 15)[53]	7,44 (49,6%)	3,86	0	15
PS (sur 10)[54]	8,39 (83,9%)	1,78	1	10

Les résultats indiquent que cette épreuve est bien réussie par les enfants de 1ère année avec 70,23% de réussite moyenne et un écart-type de 8,2.

L'analyse des performances obtenues aux différents sous-scores de l'épreuve IME1 révèle que les items OC et PS sont bien réussis par la majorité des enfants avec des taux de réussite respectifs de 81,4% (σ=5,16) et 83% (σ=1,78). Les enfants réussissent les items PLC à 67,66% (σ=3,07). En revanche, ils trouvent plus de difficultés à repérer les perturbations orthographiques portant sur les voyelles (49,6% de réussite aux items PLV et un écart-type de 3,86).

2.1.4.2. Les performances en identification de mots à T3

Comme on pouvait le prévoir, les performances des enfants à l'épreuve IME évoluent en 2ème année avec un pourcentage global de réussite d'environ 77,44% (voir tableau 10). La comparaison des scores obtenues à cette épreuve en fonction du niveau scolaire nous a permis d'observer une différence significative ($t(1,170)=77,13$; $p<.001$), ce qui témoigne d'une

[50] Identification de mots écrits à T2 (1ère année)

[51] Items congruents avec les images et correctement orthographiés.

[52] Items congruents avec les images mais ayant subis des perturbations orthographiques sur les consonnes.

[53] Items congruents avec les images mais ayant subis des perturbations orthographiques au niveau des signes vocaliques.

[54] Items non congruents avec les images (PS pour Pseudo-synonymes).

évolution des capacités d'identification de mots avec l'avancement dans la scolarité. L'épreuve Timé2 est assez bien réussie par les enfants avec un taux de réussite de 71,44%.

Tableau 10 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus aux épreuves de lecture IME2 et Timé2

Mesures	**Moyenne**	**Ecart-type**	**Minimum**	**Maximum**
IME2 (/60) [55]	46,28 (77,13%)	7,8	26	60
Timé2 (/ 36)	25,72 (71,44%)	8	4	36

La figure 8 présente les résultats obtenus par les enfants aux diffétrents sous-scores des deux épreuves d'identification de mots écrits à T3 (IME2 et Timé2). En ce qui concerne l'épreuve IME2, les meilleures performances sont obtenues avec les items OC et PS (89%). Les enfants obtiennent un taux de réussite de 78% dans l'identification des items présentant des perturbations sur les consonnes (PLC). Cependant, leurs performances dans le repérage des perturbations portant sur les voyelles restent faibles (49% de réponses correctes).

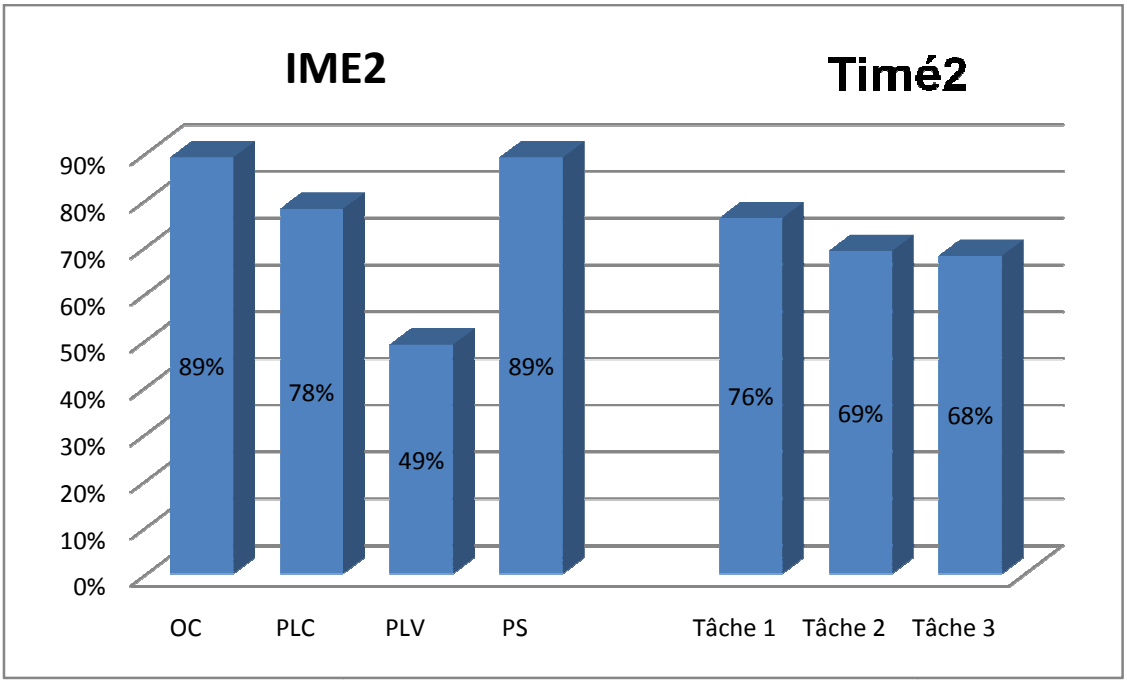


Figure 8 : performances exprimées en pourcentages moyens de réussite obtenues aux différents sous-scores des épreuves d'identification de mots à T3 (IME2 et Timé2)

[55] Identification de mots écrits à T2 (2ème année)

Les résultats obtenus aux différentes tâches de l'épreuve de lecture Timé2 montrent que les enfants obtiennent de bonnes performances dans la reconnaissance de mots fournis oralement par l'expérimentateur (Timé2.1) avec un taux de réussite de 76%. Dans la tâche de reconnaissance de mots écrits[56](Timé2.2), les enfants obtiennent un score moyen de 69% comparable à celui obtenu dans la tâche de catégorisation sémantique consistant à choisir un mot sémantiquement ou pragmatiquement proche d'un mot inducteur (Timé2.3)

2.1.4.3. Les performances en lecture-compréhension à T3

Les résultats obtenus à l'épreuve de compréhension sont présentés dans le tableau 11. Nous observons que le taux de réussite est significativement inférieur à celui obtenu à l'épreuve IME2 ($t(1,170)= 69,4$; $p<.01$) et au Timé2 ($t(1,170) =31,7$; $p<.01$). Les enfants ont réussi en moyenne 62,5% des items.

Tableau 11 : Performance moyenne de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus à l'épreuve de compréhension à T3

Mesures	Moyenne	Ecart-type	Minimum	Maximum
Cph écrit (/16)	10 (62,5%)	2,7	2	16

2.2. Analyse du lien entre les scores aux différentes épreuves

2.2.1. Analyse des corrélations

2.2.1.1. Corrélations entre les épreuves en classe préparatoire (T1)

La recherche des corrélations entre les épreuves administrées à T1 a permis de définir le degré d'association entre les différentes variables. Les résultats présentés dans le tableau 12 indiquent que les différentes mesures langagières et cognitives sont dans l'ensemble inter-corrélées de façon significative.

Nous constatons que les épreuves phonologiques : Comptage, segmentation et suppression syllabique, sont fortement reliées entre elles, ce qui montre que les capacités mobilisées dans chacune d'entre elles sont similaires. La corrélation la plus forte est celle entre le comptage et la segmentation syllabique ($r=.93$; $p<.01$). Les épreuves de conscience phonologique sont aussi corrélées positivement avec la majorité des épreuves langagières, en particulier avec la connaissance des lettres (r varie entre .38 et .54 ; $p<.05$) ; la morpho-syntaxe (r varie entre .28

[56] L'enfant devant choisir parmi cinq mots celui correspondant à une image.

et .39 ; p<.05) et la compréhension de consignes (r varie entre .31 et .43 ; p<.05) ; les corrélations les plus fortes étant obtenues avec la tâche de suppression syllabique (respectivement r=.59, r=.39 et r=.43 ; p<.05). Toutefois, les performances à l'épreuve de détection de rime entretiennent de faibles corrélations avec les différentes épreuves (r varie entre .04 et .30). Soulignons que l'épreuve de connaissance des lettres est fortement corrélée à l'épreuve morpho-syntaxique (r=.47 ; p<.01) et la compréhension de consignes (r=.44 ; p<.01). Les capacités de dénomination rapide sont, en revanche, faiblement corrélées avec la majorité des épreuves. La seule corrélation significative concerne le niveau attentionnel (r=.23 ; p<.05).

Tableau 12 : Matrice des corrélations entre les compétences langagières et cognitives évaluées en CP

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**13**	**14**	**15**	**16**	**17**
1. Déno rap[57]	1															
2. Compt syl[58]	.18	1														
3. Seg syl[59]	.12	**.93****	1													
4. Supp syl[60]	.08	**.50****	**.54****	1												
5. Intrus syl[61]	.06	.29*	.30*	**.45****	1											
6. Intrus rime[62]	.04	.29*	.30*	.24*	.26*	1										
7. Morpho-syn[63]	.09	.28*	.32*	.39*	.33*	.20*	1									
8. Conn lettre[64]	.15	.39*	**.44****	**.59****	.38*	.13	**.47****	1								
9. Cph orale[65]	.04	.12	.18	.27*	.23*	.24*	.33*	.17	1							
10. Cph consi[66]	.07	.31*	.33*	**.43****	.35*	.14	**.44****	**.44****	.33*	1						
12. Inteligence	.13	.22*	.22*	.32*	.25*	.19	**.42****	.31*	.29*	.26*	1					
13. MCT	.14	**.42***	**.46****	.36*	.31*	.26*	.36*	.35*	.25*	.32*	.26*	1				
15. Graphisme	.15	.32*	.37*	.38*	.33*	.15	**.43****	**.44****	.28*	**.41****	.34*	.34*	.05	1		
16. Concepts[67]	.07	.36*	**.43****	**.42****	.29*	.21*	.36*	.39*	.32*	.35*	.26*	.39*	.16	**.42****	1	
17. Attention	.23*	.18	.21*	.23*	.10	.07	.30*	.32*	.08	.16	.20*	.18	.00	.39*	.17	1

** s à p<.01 * s à p<.05

En gras : les corrélations les plus fortes.

En ce qui concerne les variables cognitives, on observe qu'elles sont inter-corrélées (sur les 15 corrélations, 8 sont significatives à p<.01). En particulier, les performances aux épreuves

[57] Dénomination rapide
[58] Comptage syllabique
[59] Segmentation syllabique
[60] Suppression syllabqiue
[61] Détection de l'intrus syllabique
[62] Détection de l'intrus rimique
[63] Morpho-syntaxe
[64] Connaissance des lettres
[65] Compréhension orale
[66] Compréhension de consignes
[67] Concepts de base

de graphisme et de concepts de base sont les plus corrélées (r=.42 ; p<.01). Ces dernières sont également corrélées à l'épreuve de MCT (respectivement r=.34 et r=.39 ; p<.05). Par ailleurs, les scores obtenus aux épreuves cognitives sont significativement corrélées avec les scores de certaines épreuves langagières : segmentation et suppression syllabique, morpho-syntaxe et connaissance des lettres. Les corrélations les plus importantes sont celles observées entre la MCT et les tâches phonologiques, notamment la segmentation syllabique (r=.48 ; p<.01) et entre le graphisme et la connaissance des lettres (r=.44 ; p<.01). De manière générale, les corrélations sont plus faibles que pour les variables langagières.

2.2.1.2. Corrélations entre les épreuves en 1ère année (T2)

Les analyses de corrélation réalisées entre les épreuves à T2 (tableau 13) montrent que les deux mesures de conscience phonémique (détection de l'intrus de phonème initial et de phonème final) sont significativement corrélées entre elles (r=.48 ; p<.01).

Tableau 13 : corrélations entre les différentes mesures à T2

	1	**2**	**3**
1. phon ini	1		
2. phon fin	.48**	1	
3. jug morph	.25*	.21*	1

** s à p<.01 * s à p<.05

Ces analyses montrent également des corrélations faibles mais significatives entre la tâche de jugement des liens morphologiques et les deux tâches de détection de phonèmes (initial : r=.25 ; p<.05 ; final : r=.21 ; p<.05).

2.2.1.3. Corrélations entre les épreuves à T1 et T2

L'analyse des corrélations entre les mesures à T1 et à T2 (tableau 14) montre que l'épreuve de jugement de liens morphologiques présente les coefficients de corrélation les plus forts avec les mesures à T1 dont 4 sont supérieurs ou égales à .40. Cette épreuve est fortement corrélée aux épreuves phonologiques et plus particulièrement aux épreuves de comptage, segmentation et suppression syllabique (respectivement r=.33, r=.35 et r=.42 ; p<.05), et également aux épreuves de connaissance des lettres et de compréhension de consignes (respectivement r=.40, r=.41; p<.01). Parmi les compétences cognitives, c'est la MCT qui présente la corrélation la plus forte avec le jugement de liens morphologiques (r=.40 ; p<.05).

Tableau 14 : Corrélations entre les différentes mesures à T1 et T2

Mesures à T1	**Mesures à T2**		
	Détec phonè init	**Détec phonè fin**	**Jug lien morph**
Comptage syllabique	.18	.11	.33*
Segmentation syllabique	.19	.11	.35*
Suppression syllabique	**.39***	.36*	**.42***
Détection Intrus syllabique	**.39***	**.45****	.23*
Intrus intrus rimique	.17	.16	.20*
Morpho-syntaxe	.32*	.30*	.33
Connaissance des lettres	**.41****	.30*	**.40****
Compréhension orale	.29	.15*	.24
Compréhension consignes	.30*	.25*	**.41****
Dénomination rapide	.02	.04	.15
Intelligence	.28*	.32*	.26*
MCT	.27*	.29*	**.40****
Graphisme	.34*	.22*	.25*
Concepts de base	.35*	.30*	**.38****
Attention	.15	.09	.35*

** s à p<.01 *s à p<.05

En gras : les corrélations les plus fortes

Les connaissances phonémiques observées à T2 sont fortement corrélées aux connaissances syllabiques évaluées à T1, en particulier aux tâches de suppression syllabique et de détection de l'intrus syllabique (r varie entre .36 et .45 ; p<.01). En revanche, nous n'observons pas de liens significatifs avec le comptage et la segmentation syllabique, ni avec la détection de l'intrus rimique. La corrélation la plus importante (r=.45 ; p<.01) est celle observées entre la tâche de détection de l'intrus du phonème final et la tâche de détection de l'intrus syllabique. Les connaissances phonémiques à T2 sont liées également aux compétences cognitives mesurées à T1, à l'exception du niveau attentionnel.

2.2.1.4. Corrélation entre les épreuves de lecture à T2 et T3

Les résultats présentés dans le tableau 15 indiquent que les corrélations entre les différentes épreuves de lecture utilisées en 1ère et 2ème année primaire sont fortement significatives. Ainsi, les analyses montrent une corrélation forte et significative entre les scores à l'épreuve IME à T2 (IME1) et T3 (IME2) (r=.64 ; p<.001). De la même façon, les performances au Timé2 sont fortement corrélées à l'épreuve IME (IME1 : r=.61; p<.001 et IME2 : r=.67 ; p<.001). Enfin,

les analyses confirment également l'existence d'un lien significatif entre la compréhension et l'identification de mots à T2 et T3 (IME1 : r=.51 ; p<.001 ; IME2 : r=.52 ; p<.001 ; Timé2 : r=.67 ; p<.001).

Tableau 15 : Corrélations entre les différentes épreuves de lecture à T1 et T2

	1	2	3	4
1. IME1	1			
2. IME2	.64*	1		
3. Timé 2	.61*	.67*	1	
4. cph écrit	.51*	.52*	.67*	1

* s à p<.001

2.2.2. Analyse factorielle en composantes principales (ACP)

L'objectif principal de cette ACP consiste à structurer les compétences que l'on a mesurées et à les résumer en un plus petit nombre de variables hypothétiques appelées facteurs[68]. L'ACP a été conduite sur les différentes mesures à T1, T2 et T3 séparément.

2.2.2.1. Analyse en composantes principales sur les mesures à T1

L'ACP réalisée avec une rotation de type oblimin a permis de regrouper l'ensemble des épreuves langagières et cognitives de CP en 3 facteurs. Les contributions factorielles de chaque composante sont présentées dans le tableau 16. Afin de vérifier la fiabilité de cohérence interne de chaque composante, nous avons eu recours au coefficient « Alpha de Cronbach » (α de Cronbach)[69]. L'ACP a permis de regrouper l'ensemble des épreuves langagières et cognitives de CP en 3 facteurs. Cette solution factorielle explique 52,70% de la variance totale :

- Le premier facteur, expliquant à lui seul 22,66% de la variance, regroupe les épreuves qui nécessitent une manipulation des unités phonologiques : comptage (compt syl), segmentation (Seg syl), suppression syllabique (Supp syl), détection de l'intrus syllabique, l'épreuve de connaissance des lettres et la MCT.

[68] L'ACP consiste à réduire les variables initiales en un nombre réduits de facteurs. Ces derniers résultent de combinaisons linéaires des variables initiales dont les coefficients sont fournis par les poids factoriels. Selon la valeur de ces coefficients, il apparaît que certaines variables initiales contribuent plus que d'autres à la formation d'un axe particulier.

[69] Un alpha de Cronbach est acceptable quand il est supérieur à .60.

- Le second facteur qui représente 20,5% de la variance est saturé exclusivement sur les épreuves morpho-flexionnelles : morphosyntaxe et jugement de l'ordre, et l'épreuve de compréhension orale.

- Le troisième facteur, dont la contribution à la variance est de 9,54%, associe trois tâches cognitives : intelligence, graphisme et concepts de base.

Tableau 16 : Matrice de l'analyse factorielle en composantes principales des mesures à T1 (les saturations inférieures à 0.5 ne sont pas indiquées)

	Facteur 1	**Facteur 2**	**Facteur 3**
Compt syl	.891		
Seg syl	.915		
Supp syl	.744		
Détec intrus rime			
Détec intrus syl			
Conn lettres	.622		
Jugement ordre		.713	
Morpho-syntaxe		.725	
Cph orale		.646	
Cph consignes			
Déno rap			
Graphisme			.573
Intelligence			.655
Concepts base			.528
MCT	.528		
Attention			
% de variance	**22,66**	**20,5**	**9,54**
α de Cronbach	**.82**	**.64**	**.55**

On constate une différenciation assez nette entre les dimensions langagières (saturées exclusivement dans les facteurs 1 et 2) et cognitives (saturées exclusivement dans le facteur 3). Ce sont les facteurs 1 et 2 qui contribuent le plus à la variance (respectivement 22% et 20%). Les tests de fiabilité de cohérence interne indiquent que chacune des composantes présente une bonne homogénéité interne (valeur des alphas de Cronbach supérieures à .60), excepté pour la dimension cognitive.

L'analyse des corrélations entre les composantes (tableau 17) indique que la relation entre les deux premiers facteurs est forte et significative (r=.45). Ces deux facteurs partagent effectivement une dimension en commun. Ils regroupent des épreuves qui nécessitent de la part du sujet d'identifier les informations pertinentes pour procéder à diverses opérations d'analyse et de manipulation du langage oral. Cependant, les facteurs 1 et 2 ne semblent pas fortement corrélés au facteur 3 (respectivement r=.21 et r=.19).

Tableau 17 : Corrélations entre les différents facteurs à T1

	Facteur 1	**Facteur 2**	**Facteur 3**
Facteur 1	1		
Facteur 2	**.45**	1	
Facteur 3	.21	.19	1

2.2.2.2. Analyse en composantes principales sur les mesures à T2

Une ACP conduite sur les mesures à T2 permet de dégager un seul facteur qui rend compte de 55,16% de la variance (tableau 18). Ce facteur associe les deux épreuves de détection de phonème initial et final et l'épreuve de jugement des liens morphologiques.

Tableau 18: Matrice de l'analyse factorielle en composantes principales des mesures à T2

	Facteur 1
Phonè fin	.800
Phonè ini	.823
Jug morph	.581
% de variance	**55,16**
α de Cronbach	**.42**

L'estimation de la cohérence interne à l'aide de l'Alpha de Cronbach révèle un coefficient de .42. Ce faible coefficient peut être expliqué par le nombre réduit d'épreuves prises en considération dans cette analyse. En effet, plus le nombre d'items est important, plus la valeur de l'alpha de Cronbach est élevée.

2.2.2.3. Analyse en composantes principales sur les mesures à T3

L'ACP menée sur les mesures à T3 permet de mettre en évidence un facteur dont la contribution à la variance est de 76,97 % (tableau 19). Il regroupe les épreuves de lecture : identification de mots écrits en en 2ème année (IME2) (poids de .87), Timé2 (poids de .93), et compréhension écrite (Cph écrit) (poids de .82). Ce facteur correspond aux compétences d'identification de mots et de compréhension.

Tableau 19: Matrice de l'analyse factorielle en composantes principales des mesures à T3

	Facteur 1
IME2	.872
Timé2	.932
Cph écrit	.824
% de variance	**76,97**
α de Cronbach	**.82**

Le coefficient α de Cronbach égal à .82 correspond à un niveau satisfaisant de fiabilité de cohérence interne.

2.3. Synthèse

L'examen des performances obtenues aux différentes évaluations langagières et cognitives effectuées à T1, T2 et T3 mettent en évidence les points suivants :

- Les enfants tunisiens prélecteurs ont un certain degré de conscience phonologique, en particulier une conscience des unités syllabiques. Il semble que la capacité à identifier et à manipuler les unités larges ne leur pose pas de problème avant le début de l'apprentissage de la lecture. Toutefois, les scores planchers obtenus aux épreuves phonémiques montrent que la conscience phonémique ne se développe vraiment qu'avec l'apprentissage de la lecture. Les élèves de CP semblent incapables de décomposer les mots en unités plus petites que la syllabe. Les résultats aux épreuves évaluant les connaissances morphologiques suggèrent que les enfants d'âge préscolaire maîtrisent les règles flexionnelles (52% de réussite à l'épreuve morpho-syntaxique). En revanche, nous pouvons constater que les connaissances morphologiques dérivationnelles évaluées par une épreuve de jugement de liens morphologiques sont très faibles (effet plancher) indiquant que les enfants de CP ne maîtrisent pas encore la notion de famille morphologique. Cependant, ces connaissances se développent assez bien en 1ère année (moyenne de 65% de réussite).
- Les analyses des corrélations à T1 montrent que les épreuves langagières et plus particulièrement celles mesurant les compétences phonologiques sont fortement corrélées entre elles. Ces compétences sont étroitement liées aux tâches morphologiques et à la connaissance des lettres.

  Des liaisons fortes et significatives ont été observées entre les épreuves langagières et certaines épreuves cognitives. En effet, la MCT apparaît fortement impliquée dans les épreuves de conscience phonologique. De même, le niveau de graphisme apparaît très lié

à la connaissance des lettres.

Les analyses de corrélation à T2 indiquent une liaison étroite entre les connaissances phonémiques et les connaissances morphologiques dérivationnelles.

Les analyses de corrélation entre les épreuves de lecture indiquent une liaison étroite d'une part entre les différentes épreuves d'identification de mots écrits à T2 et T3 et d'autre part entre les capacités d'identification de mots et de compréhension.

- Une analyse en composantes principales (ACP) effectuées sur les différentes mesures à T1, T2 et T3 a permis de regrouper les variables initiales en un nombre réduit de facteurs. L'ACP réalisées sur l'ensemble des épreuves à T1 permet de dégager 3 facteurs. Le premier regroupe les épreuves qui nécessitent une manipulation des unités phonologiques qui apparaissent avoir le plus de poids dans la construction de ce facteur. La connaissance des lettres et la MCT se trouvent associées à ces épreuves, ce qui laisse supposer que ces mesures sont des bons indicateurs des compétences phonologiques des enfants.

  L'épreuve qui a le plus de poids dans la construction de ce facteur est l'épreuve de segmentation syllabique. Ce facteur permet ainsi de mesurer les « connaissances grapho-phonologiques ».

  Le second facteur regroupe les épreuves morpho-syntaxiques qui ont le plus de poids dans la construction de ce facteur avec l'épreuve de compréhension orale. Ce facteur évalue notamment l'aspect morphosyntaxique qui est très important en lecture puisqu'il sous-tend la structure des textes et guide l'interprétation que l'on peut en faire. Ce facteur peut être assimilé à des « connaissances morpho-flexionnelles ».

  Le troisième facteur associe les épreuves qui mesurent « les compétences cognitives » : niveau intellectuel, concepts de base et graphisme.

  L'ACP réalisée sur les épreuves à T2 permet de mettre en évidence un facteur composé des épreuves phonémiques et de l'épreuve de jugement des liens morphologiques. Mais ce sont surtout les tâches phonémiques qui interviennent avec un poids factoriel élevé dans la constitution de ce facteur qui peut être appelé « connaissances phonémiques ».

  En 2ème année, l'ACP permet de mettre en évidence un facteur composé des épreuves d'identification de mots et de compréhension. Ce facteur permet de mesurer les « performances en lecture ».

2.4. Les relations prédictives entre les compétences langagières et cognitives et les épreuves de lecture

L'objectif de cette partie est de déterminer les relations prédictives entre les variables mesurées en CP et en 1ère année et le niveau ultérieur en lecture.

2.4.1. Prédicteurs de l'apprentissage de la lecture en fin de 1ère année primaire (T2)

Une première série d'analyse examine les corrélations entre les compétences cognitives et langagières précoces des enfants et l'efficience ultérieure en lecture évaluée par une épreuve d'identification de mots écrits (IME1). Dans un second temps, on a cherché, au moyen d'analyses de régressions multiples (stepwise regression), à déterminer les meilleurs prédicteurs des capacités ultérieures en identification de mots écrits.

2.4.1.1. Analyse des corrélations entre les épreuves à T1 et le niveau d'identification de mots à T2

Le tableau 20 présente les corrélations entre les différentes compétences mesurées en CP et le niveau de lecture en 1ère année évalué par l'épreuve d'identification de mots écrits (IME1). Nous constatons que quatre variables langagières sont corrélées positivement à plus de .40 à l'épreuve IME1. Les corrélations les plus élevées concernent la connaissance des lettres (r=.47 ; p<.01), la suppression syllabique (r=.46 ; P<.01), la morpho-syntaxe (r=.45 ; P<.01) et la compréhension de consignes (r=.44 ; P<.01) ; viennent ensuite la mémoire à court terme (r=.38 ; P<.05), la segmentation syllabique (r=.37 ; P<.05), le graphisme et le niveau intellectuel (r=.34 ; P<.05). Inversement, la détection de l'intrus rimique, le niveau d'attention et la capacité de dénomination rapide n'entretiennent pas de corrélations significatives avec l'épreuve d'identification de mots.

Il apparait ainsi que les évaluations langagières et cognitives réalisées en CP sont liées aux performances en lecture en 1ère année. Toutefois, les corrélations les plus fortes sont obtenues avec les épreuves langagières. Parmi les compétences cognitives, la MCT semble la plus corrélée à l'identification de mots (r=.38 ; P<.05).

Tableau 20 : Corrélations entre les compétences cognitives et langagières en CP et les performances en lecture en 1ère année (IME1 : identification de mots écrits à T2)

	IME1
Comptage syllabique	.32*
Segmentation syllabique	**.37***
Suppression syllabique	**.46****
Intrus syllabique	.28*
Intrus rime	.19ns
Morpho-syntaxe	**.45****
Connaissance des lettres	**.47****
Compréhension orale	.22*
Compréhension consignes	**.44****
Dénomination rapide	.10ns
Intelligence	.34*
MCT	**.38***
Graphisme	**.**34*
Concepts de base	.31*
Attention	.17ns

** s à p<.01 * s à p<.05

En gras : les corrélations les plus fortes.

2.4.1.2. Analyse des corrélations entre les épreuves à T2 et le niveau d'identification de mots à T2

L'analyse des corrélations entre les épreuves à T2 et le niveau d'identification de mots à T2 (tableau 21) montre que les scores obtenus à la tâche de détection de l'intrus phonémique est corrélée significativement aux performances d'identification de mots de la même année (initial : r=.34 ; p<.05 ; final : r=.27 ; p<.05). De la même façon, les performances obtenues à la tâche de jugement des liens morphologiques sont fortement corrélées à l'épreuve IME1 (r=.46 ; p<.01).

Tableau 21 : Corrélations entre les compétences à T2et les performances en lecture en 1ère année (IME1 : identification de mots écrits à T2)

	IME1
Jug liens morph	**.46****
Détect phonè initial	.34*
Détection phonè final	.27*

** s à p<.01 * s à p<.05

En gras : la corrélation la plus forte.

2.4.1.3. Analyse de régressions pas à pas sur la variable « identification de mots écrits à T2 »

Nous avons cherché, au moyen de régressions pas à pas, à déterminer les variables les plus prédictives des performances en lecture en 1ère année en accordant aux scores des élèves à l'épreuve d'identification de mots écrits à T2 (IME1) le statut de variable dépendante. Les mesures cognitives et langagières à T1 ont été considérées comme des variables explicatives.

Tableau 22: Analyse de régression pas à pas entre les mesures à T1 et les performances en identification de mots à T2

Variables Explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Pas 1 : Connaissance des lettres	.241	.254	.245	56.152
Pas 2 : Suppression syllabique	.274	.321	.312	39.935
Pas 3 : Morpho-syntaxe	.233	.359	.346	30.917

Les résultats présentés dans le tableau 22 indiquent que les variables contribuant le plus à expliquer le niveau d'identification de mots sont celles mesurées par les épreuves de connaissance des lettres, de suppression syllabique et de morpho-syntaxe. La connaissance des lettres apporte la contribution la plus importante (24%). La suppression syllabique, incluse dans un deuxième temps, apporte un gain supplémentaire de pouvoir explicatif d'environ 7%. La morpho-syntaxe ajoute 3% de variance supplémentaire.

Lorsqu'on examine les résultats en prenant pour variables explicatives les mesures à T2 (voir tableau 23), nous constatons que la variable « jugement de liens morphologiques » apporte la contribution la plus importante (21%). En second lieu, la variable « détection du phonème initial» est introduite dans le modèle occupant ainsi la seconde place avec 7% de variance expliquée.

Tableau 23 : Analyse de régression pas à pas entre les mesures à T2 et les performances en identification de mots à T2

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Pas 1 : Jug liens morph	.403	.223	.218	48.425
Pas 2 : Détect phonè initial	.270	.291	.282	34.458

2.4.2. Prédicteurs de l'apprentissage de la lecture en fin de 2ème année primaire (T3)

En 2ème année, les enfants ont passé deux nouvelles épreuves de lecture, en plus de celle

utilisée en 1ère année : un test d'identification de mots écrits inspiré du Timé2 (Ecalle, 2003) et une épreuve de compréhension écrite. Dans cette partie, nous allons examiner les relations prédictives entre les variables mesurées en CP (T1) et en 1ère année (T2) et les performances en lecture en 2ème année (T3).

2.4.2.1. Analyse des corrélations

a. Analyse des corrélations entre les mesures à T1 et le niveau d'identification de mots à T3

Le tableau 24 représente les coefficients de corrélation entre les mesures à T1 et les performances en identification de mots à T3 (IME2).

Tableau 24 : Corrélations entre les mesures à T0 et les performances en identification de mots à T2

	IME2	**Timé2**
Comptage syllabique	.35*	.35*
Segmentation syllabique	**.41****	**.42***
Suppression syllabique	.37*	**.41****
Intrus syllabique	**.38***	.36*
Intrus rime	.17ns	.16 ns
Morpho-syntaxe	.32*	**.38****
Connaissance des lettres	**.43****	**.51****
Compréhension orale	.18 ns	.31*
Compréhension consignes	.37*	**.46****
Dénomination rapide	.19ns	.14 ns
Intelligence	.31*	.35*
MCT	.28*	.36*
Graphisme	.36*	**.40****
Concepts de base	.29*	.37*
Attention	.20*	.37*

** s à p<.01 * s à p<.05

En gras : les corrélations les plus fortes.

Nous relevons des relations significatives entre la plupart des variables cognitives et langagières et les différentes épreuves de lecture. Les plus fortes corrélations relevées à l'épreuve IME2 concernent la connaissance des lettres (r=.43 ; p<.01), la segmentation syllabique (r=.41 ; p<.01) et la détection de l'intrus syllabique (r=.38 ; p<.05). Les corrélations observées avec le Timé2 sont plus élevées puisque 6 corrélations sont supérieures

ou égales à .40. Les corrélations les plus fortes concernent la connaissance des lettres (r=.51 ; p<.01), la compréhension de consignes (r=.46 ; p<.01) et la segmentation syllabique (r=.42 ; p<.01).

Les deux matrices de corrélations présentent de fortes similitudes. Toutefois, les corrélations obtenues avec le Timé2 sont globalement plus élevées que celles obtenues avec l'épreuve IME (corrélations moyennes = .35 pour Timé2 *vs* .30 pour l'épreuve IME2).

b. Analyse de corrélations entre les mesures à T2 et le niveau d'identification de mots à T3

Parmi les épreuves utilisées en 1ère année (tableau 25), c'est l'épreuve de jugement de liens morphologiques qui est la plus corrélée à l'épreuve IME2 (r=.37 ; p<.05) et au Timé2 (r=.46 ; p<.01). Nous constatons également une corrélation significative entre la tâche de détection du phonème initial et les deux épreuves d'identification de mots (Timé2 : r=.39 ; p<.05 et IME2 : r=.29 ; p<.05). La corrélation entre la détection du phonème final et les deux épreuves de lecture est assez faible avec un résultat meilleur au Timé2 (IME2 : r=.20 ; p<.05, Timé2 : r=.28 ; p<.05).

Tableau 25 : Corrélations entre les mesures à T2et les performances en identification de mots à T3

	IME2	**Timé2**
Détect phonè initial	.29*	**.39***
Détect phonè final	.20*	.28*
Jug liens morph	**.37***	**.46****

** s à p<.01 * s à p<.05 En gras : les corrélations les plus fortes.

c. Analyse des corrélations entre les mesures à T1 et le niveau de compréhension à T3

Le tableau 26 présente les corrélations entre les différentes épreuves langagières et cognitives à T1 et le niveau de compréhension écrite à T3. Nous constatons que les relations entre les différentes mesures de CP et les performances à l'épreuve de compréhension écrite sont moins fortes que celles observées avec les épreuves IME2 et Timé2. Les liaisons les plus fortes sont obtenues avec la morpho-syntaxe (r=.41 ; p<.01), la compréhension de consignes (r=.40 ; p<.01) et la détection de l'intrus syllabique (r=.38 ; p<.05).

Tableau 26 : Corrélations entre les mesures à T1et les performances en compréhension écrite à T3

	Cph écrit
Comptage syllabique	.22*
Segmentation syllabique	.26*
Suppression syllabique	**.**34*
Intrus syllabique	**.38***
Intrus rime	.20*
Morpho-syntaxe	**.41****
Connaissance des lettres	.33*
Compréhension orale	.23*
Compréhension consignes	**.40****
Dénomination rapide	.04ns
Intelligence	.28*
MCT	.31*
Graphisme	.37*
Concepts de base	.27*
Attention	.26*

** sig à p<.01 * sig à p<.05

En gras : les corrélations les plus fortes

d. Analyse des corrélations entre les mesures à T2 et le niveau de compréhension à T3

Les résultats présentés dans le tableau 27 montrent que les compétences phonémiques et morphologiques mesurées à T2 sont significativement corrélées à la compréhension écrite à T3. Les corrélations obtenues pour les trois épreuves sont presque de même niveau.

Tableau 27 : Corrélations entre les mesures à T2 et les performances en lecture-compréhension à T3

	Cph écrit
Détect Phonè initial	.33*
Détect Phonè final	.32*
Jug liens morph	.34*

* sig à p<.05

2.4.2.2. Analyse de régressions multiples

A l'aide d'analyses de régression pas à pas, nous avons cherché à mettre en évidence la valeur prédictive et la contribution des compétences langagières et cognitive à l'explication du niveau de lecture en 2ème année primaire.

a. Analyse de régressions sur l'épreuve IME à T3 (IME2)

En 2ème année, l'analyse de régression pas à pas réalisée entre les mesures cognitives et langagières à T1 et les performances à l'épreuve d'identification de mots écrits à T3 (tableau 28) montre que la connaissance des lettres occupe la première position avec 18% de variance expliquée. La segmentation syllabique apporte un gain supplémentaire de pouvoir explicatif d'environ 6%. Enfin, la détection de l'intrus syllabique contribue assez modestement à la variance totale expliquée (3% de variance supplémentaire). Les trois variables retenues expliquent 27% de la variance en identification de mots en 2ème année.

Tableau 28 : Analyse de régression pas à pas entre les mesures à T1 et les performances en identification de mots à T3

Variables explicatives	**β**	**R^2**	**R^2 modifié**	**F**
Pas 1: Conn des lettres	.265	.187	.182	38.820
Pas 2: Seg syllabique	.245	.249	.240	27.828
Pas 3: Détect intrus syllabique	.159	.270	.275	20.582

Dans une seconde analyse de régression, les performances à l'épreuve de jugement de liens morphologiques et les performances aux deux tâches de détection de l'intrus phonémique (qui sont les trois variables de 1ère année) ont été introduites comme une variable indépendante. On observe que les connaissances morphologiques sont extraites en première position et expliquent 13% de la variance (tableau 29). La détection du phonème initial ajoute 4% de variance supplémentaire.

Tableau 29 : Analyse de régression pas à pas entre les mesures à T2 et les performances en identification de mots à T3

Variables explicatives	**B**	**R^2**	**R^2 ajusté**	**F**
Pas 1: Jug liens morph	.314	.137	.132	26.778
Pas 2: Détect phonè initial	.219	.182	.172	18.628

b. Analyse de régressions sur l'épreuve Timé2 à T3

Nous avons conduit également une analyse de régression en prenant pour variable dépendante les scores obtenus au Timé2. Nous avons introduit dans un premier temps les mesures utilisées à T1. Les résultats présentés dans le tableau 30 confirment la contribution de la connaissance des lettres (31% de variance expliquée). La segmentation syllabique apparaît au second pas avec 8% de pouvoir explicatif. La dernière variable retenue dans le modèle est la compréhension de consignes (3% de la variance).

Tableau 30 : Analyse de régression pas à pas entre les mesures à T1 et les performances à l'épreuve Timé2 à T3

Variables explicatives	β	R^2	R^2 ajusté	F
Pas 1: Conn des lettres	.354	.263	.268	60.200
Pas 2 : Seg syllabique	.229	.341	.343	41.612
Pas 3 : Compréh de consignes	.173	.385	.373	30.592

Lorsqu'on effectue une analyse de régression en prenant comme variables indépendantes les scores aux épreuves passées à T2, deux variables émergent (tableau 31) : le jugement de liens morphologiques et la détection de l'intrus de phonème initial, dont la contribution à la variance est de 28%. Le jugement de liens morphologiques explique la plus grande part de variance des performances en lecture (21%). La détection du phonème initial ajoute une part de variance non négligeable (7% de variance supplémentaire).

Tableau 31 : Analyse de régression entre les mesures à T2 et les performances à l'épreuve timé2 à T3

Variables explicatives	β	R^2	R^2 ajusté	F
Pas 1 : Jug liens morph	.392	.218	.213	47.052
Pas 2 : Détect phonè initial	.292	.297	.289	35.523

c. Analyse de régressions sur l'épreuve de compréhension écrite à T3

Les résultats présentés dans le tableau 32 montrent que les performances en compréhension écrite sont principalement expliquées par la morpho-syntaxe, la détection de syllabe initiale et la compréhension de consignes. Ces trois variables expliquent 28% de la variance totale. La morpho-syntaxe permet d'expliquer de façon significative 16% de la variance en compréhension. La détection de l'intrus syllabique, incluse dans un deuxième temps, apporte

un gain de pouvoir explicatif de 10%. Enfin, la compréhension de consignes ajoute 3% de variance supplémentaire.

Tableau 32 : Analyse de régression entre les mesures à T1 et les performances à l'épreuve de lecture-compréhension à T3

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Pas 1: Morpho-syntaxe	.221	.161	.156	32.529
Pas 2 : Detect Syllabe initiale	.257	.270	.257	25.689
Pas 3 : Compréh de consignes	.213	,293	.289	20.620

Lorsqu'on effectue une seconde analyse de régression en prenant les mesures de 1ère année comme des variables indépendantes (tableau 33), nous constatons que la variable « jugement des liens morphologiques » est extraite en 1ère position avec 11% de variance. La détection du phonème final et initial apporte un gain de pouvoir explicatif d'environ 8%.

Tableau 33 : Analyse de régression entre les mesures à T2 et les performances à l'épreuve de lecture-compréhension à T3

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Pas 1 : Jug liens morph	.260	.118	.113	22.642
Pas 2 : Détect phonè final	.179	.183	.173	18.820
Pas 3 : Détect phonè initial	.179	.207	.192	14.508

2.4.3. Synthèse sur la contribution des différentes variables à l'explication des performances en lecture

Ces résultats mettent en évidence les points suivants :

- Les évaluations langagières de CP (T1) sont assez fortement liées aux performances en lecture en 1ère année (T2) et en 2ème année du primaire (T3), quelle que soit l'épreuve de lecture considérée. Les connaissances phonologiques, développées avant le début de l'apprentissage de la lecture, sont fortement corrélées avec les performances en lecture (identification de mots et compréhension) une année et deux ans plus tard, à l'exception des connaissances rimiques. La connaissance des lettres, la morpho-syntaxe et la compréhension de consignes sont également très reliées aussi bien aux performances ultérieures en identification de mots qu'en compréhension. Les connaissances les plus associées aux capacités de compréhension sont les connaissances morpho-syntaxiques.

La lecture en 1ère année et en 2ème année est également significativement corrélée avec certaines épreuves cognitives, et plus particulièrement le graphisme, l'intelligence et la MCT.

- L'observation des corrélations entre les mesures à T2 et le niveau de lecture à T3 montre une relation forte et significative entre le jugement de liens morphologiques et les différentes épreuves de lecture à T2 et T3.

- L'analyse des corrélations entre les différentes mesures à T1 et T2 et les performances en lecture montrent une certaine variabilité en fonction de l'épreuve de lecture considérée. Les corrélations les plus fortes étant obtenues avec le score global au Timé2 (voir figure 9). Les corrélations avec l'épreuve d'identification de mots écrits sont globalement plus faibles que celles obtenues en 1ère année avec la même épreuve.

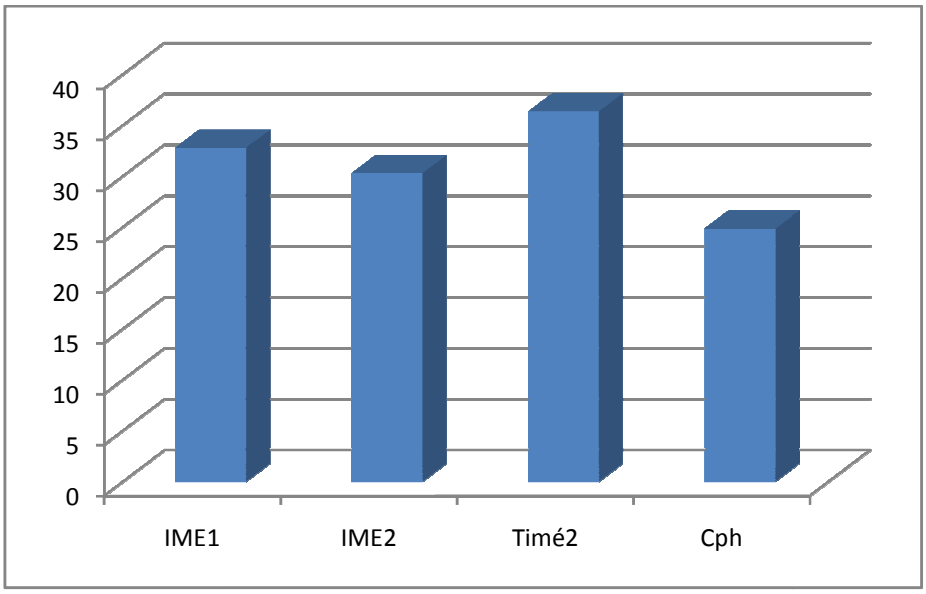


Figure 9 : Moyenne des corrélations obtenues avec les épreuves de lecture en 1ère année (T2) et en 2ème année (T3)

- De façon générale, les analyses de régression montrent que les variables prédictives mesurées à T1 et T2 apportent une contribution non négligeable à la variance en lecture, quelle que soit la variable dépendante à expliquer (tableaux 34 et 35).

Tableau 34 : Récapitulatif des meilleurs prédicteurs de l'apprentissage de la lecture en CP (T1)

Prédicteurs CP			
IME1	**IME2**	**Timé2**	**Compréhension**
Conn des lettres Supp syllabique Morpho-syntaxe	Conn des lettres Seg syllabique Détec syll initiale	Conn des lettres Seg syllabique Cph consignes	Morpho-syntaxe Détect intrus syll Cph consignes
34% de variance expliquée	**27% de variance expliquée**	**37% de variance expliquée**	**28% de variance expliquée**

Tableau 35 : Récapitulatif des meilleurs prédicteurs de l'apprentissage de la lecture en 1ère année (T2)

Prédicteurs 1ère année		
IME2	**Timé2**	**Compréhension**
Jug liens morph Détec phonè init	Jug liens morph Détec phonè init	Jug liens morph Détec phonè fin Détec phonè init
17% de variance expliquée	**28% de variance expliquée**	**19% de variance expliquée**

3. Discussion

L'objectif principal de ce chapitre était de déterminer les meilleurs prédicteurs de l'apprentissage de la lecture en langue arabe. A cette fin, nous avons réalisé une étude longitudinale prédictive consistant à évaluer à plusieurs reprises les mêmes enfants. Au cours de la première évaluation, différentes épreuves cognitives et langagières ont été administrées à des élèves en classe préparatoire (CP) lors de passations individuelles et collectives. L'efficience en lecture a été évaluée chez les mêmes enfants en 1ère année primaire par une épreuve d'identification de mots écrits en contexte imagé (IME1). Nous avons administré également aux enfants les épreuves pour lesquels nous avons eu des scores planchers en CP (jugement des liens morphologiques et détection de l'intrus phonémique). Une seconde évaluation du niveau en lecture des enfants avait été effectuée à la fin de la 2ème année primaire. Trois épreuves ont été utilisées : l'épreuve d'identification de mots utilisée en 1ère année (IME2), un test d'identification de mots écrits inspiré du Timé2 et une épreuve de lecture-compréhension. Il était donc possible de se faire une idée sur les progrès enregistrés en

lecture et de l'évolution éventuelle de la valeur prédictive des différentes mesures à l'égard de la réussite en lecture.

Performances aux différentes épreuves

L'analyse des performances en CP indique que les enfants prélecteurs manifestent une sensibilité précoce aux unités syllabiques mais ils ont encore des difficultés à décomposer les mots en unités plus petites que la syllabe. Les scores planchers obtenus à l'épreuve de détection de l'intrus phonémique mettent en évidence les difficultés rencontrées par les élèves d'âge préscolaire à opérer explicitement des manipulations sur des phonèmes. La détection précoce des syllabes est sous-tendue par le développement du langage oral alors que l'accès à la conscience phonémique parait soumis à un apprentissage systématique et n'apparaît jamais hors de cette condition (Ecalle & Magnan, 2002 ; Ecalle & Magnan, 2007a ; Gombert, 1992a ; 1992b ; Gombert & al., 2004). Ces données confortent les résultats obtenus dans le cadre des travaux entrepris en langues latines postulant que la conscience des unités larges (syllabe, attaque-rime) est plus précoce que la conscience phonémique et que le développement de la conscience phonémique va se faire graduellement (au cours des deux premières années de l'apprentissage de la lecture) (Demont & Gombert, 2007 ; Gombert, 1990 ; 1992a ; Ecalle & al., 2002). Ces résultats sont en accord avec ceux obtenus par Ammar (1997) et Boukadida (2008) mettant en évidence la difficulté que présente l'acquisition de l'analyse phonémique par rapport à l'analyse syllabique en langue arabe.
En revanche, nos résultats pour la tâche de détection de rimes ne sont pas en accord avec ceux des études entreprises en langue anglaise (Goswami & Bryant, 1990) mais confortent les résultats obtenus en espagnol (Defior-Citoler & Torres, 2003) qui suggèrent que ces tâches demandent un plus haut niveau de conscience phonologique puisqu'elles se réfèrent à des unités infra-syllabiques. Le développement de la conscience phonologique semble donc affecté par les caractéristiques propres de la langue. Ammar (2003) trouve, en effet, que les enfants arabophones ne seraient pas sensibles à la rime puisque la structure bidimensionnelle du mot, formée d'une racine consonantique et d'un schème vocalique, serait incompatible avec le découpage attaque/rime.

Les résultats montrent également que le degré de réussite dans les tâches phonologiques dépend non seulement du son à manipuler mais aussi de la tâche à réaliser. Ainsi, la tâche de suppression est moins réussie que celles de segmentation et de comptage. Conformément aux données rapportées par Gombert (1990), cette tâche semble nécessiter un niveau de

compétence phonologique différent. Les opérations cognitives exigées lors de la manipulation d'unités phonologiques sont variables selon le niveau de difficulté de la tâche proposée. Gombert, Gaux & Demont (1994) ont trouvé, de leur côté, que le pourcentage de réussite des enfants de GSM[70] à des tâches de suppression de la syllabe finale et initiale est de 31,7%.

Par ailleurs, les analyses révèlent qu'en l'absence d'un enseignement ciblé sur l'apprentissage des lettres, les enfants développent des connaissances dans ce domaine avant l'apprentissage formel de la lecture. Ces résultats sont en accord avec ceux obtenus par Ecalle & al. (2009 ; 2008). En Tunisie, la connaissance du nom des lettres est apprise à l'heure actuelle en classe préparatoire. A leur entrée en 1ère année, les enseignants considèrent souvent que ce travail n'est plus à faire et commencent leur enseignement de la lecture en considérant des groupes de lettres ou directement des mots.

En CP, les enfants maîtrisent les règles morpho-flexionnelles et sont capables de faire un jugement sur la structure morpho-syntaxique des phrases. En revanche, à ce niveau scolaire, ils ne possèdent pas de connaissances morpho-dérivationnelles puisqu'ils ont obtenu des scores qui ne se distinguent pas du niveau de hasard à l'épreuve de jugement de liens morphologiques. Une année plus tard, nous avons noté un développement de ces connaissances puisque les enfants commencent à intégrer la notion de famille morphologique avec un taux de réussite qui se situe autour de 65%. Ces résultats confirment que l'acquisition des règles flexionnelles est plus facile que celle des règles dérivationnelles quelle que soit la langue. Ces données sont en accord avec celles issues des recherches menées sur le français (Casalis & Louis-Alexandre, 2000), l'anglais (Carlisle & Namanbhoy, 1993) et l'hébreu (Ravid & Malenky, 2001 ; cité par Boukadida, 2008) qui ont trouvé, chez des enfants d'âge préscolaire, que les structures flexionnelles sont mieux distinguées et mieux manipulées que les structures dérivationnelles. La maîtrise des règles dérivationnelles s'acquiert plus tardivement et plus lentement que celles des règles flexionnelles.

En ce qui concerne les épreuves cognitives, nous avons constaté qu'une forte majorité des élèves de CP maîtrisent les notions de temps et d'espace (81% de bonnes réponses) et ont des capacités graphiques non négligeables (68% de réponses correctes).

[70] La classe préparatoire dans le système éducatif tunisien

Analyse des liens entre les différentes épreuves

Les analyses de corrélation entre les différentes épreuves en CP (T1) montrent que les épreuves langagières sont fortement corrélées entre elles. L'existence de corrélations significatives entre les différentes tâches phonologiques montre que les capacités mobilisées par chacune d'entre elles sont similaires. Ces compétences sont étroitement liées à la connaissance des lettres et aux compétences morphologiques. Ces résultats corroborent les travaux récents qui ont établi un lien étroit entre la connaissance du nom des lettres et la conscience phonologique (Ecalle & al., 2009 ; Foulin, 2005 ; Foulin & Pacton, 2006) et entre conscience phonologique et conscience morphologique (Boukadida, 2008 ; Carlisle, 1995 ; Carlisle & Nomanbhoy, 1993 ; Casalis & Colé, 2009). Une sensibilité phonologique précoce combinée à la connaissance du nom des lettres est nécessaire pour permettre à l'enfant d'accéder au principe alphabétique. En effet, les enfants commencent à relier l'oral et l'écrit sur la base du nom des lettres (Biot-chevrier & al., 2008 ; Ecalle & al., 2009). Le niveau d'analyse phonologique contribue également à la réussite aux épreuves de morphologie (Carlisle & Nomanbhoy, 1993).

Des liaisons fortes et significatives ont été observées entre les épreuves langagières et certaines épreuves cognitives. En effet, les analyses de corrélation et l'analyse en composantes principales montrent que la conscience phonologique repose sur le développement d'autres fonctions cognitives, notamment la mémoire à court terme. Les épreuves de conscience phonologique explorent tout autant les capacités d'analyse segmentale que l'aptitude à maintenir l'information en mémoire (Demont, 2002). L'analyse explicite des unités phonologiques nécessite le maintien dans la mémoire phonologique des items à comparer (détection de l'intrus phonologique) ou de l'item à segmenter (segmentation et suppression phonologique). Ainsi, les difficultés à réaliser des épreuves de conscience phonologiques seraient susceptibles d'être liées à des difficultés de maintien et/ou d'utilisation des codes phonologiques au niveau de la boucle articulatoire (Demont, 2003).

L'analyse des corrélations à T2 montrent des liaisons fortes entre l'épreuve de jugement des liens morphologiques et les épreuves de détection de l'intrus phonémique et confirment l'existence d'un lien entre conscience phonologique et analyse morphologique (Boukadida, 2008 ; Carlisle & Nomanbhoy, 1993 ; Lecoq & al., 1996 ; Marec-Breton, 2003). Les analyses de corrélation entre les épreuves de lecture indiquent une liaison étroite entre les capacités d'identification de mots et de compréhension. En effet, le but de l'apprentissage de la lecture

consiste à développer des procédures automatiques de reconnaissance des mots écrits qui permettront d'atteindre un niveau de compréhension de ce qui est lu, égal à celui qu'il est capable d'atteindre à l'oral (Colé & Fayol, 2000).

Prédicteurs de l'apprentissage de la lecture en 1ère année (T2)

Les corrélations entre les mesures en CP (T1) et la lecture en 1ère année (T2) montrent que les variables langagières sont plus fortement corrélées aux scores ultérieurs en lecture que les variables cognitives. Les connaissances phonologiques, développées avant le début de l'apprentissage de la lecture, sont corrélées avec les performances en lecture à la fin de la 1ère année, à l'exception des connaissances rimiques. Conformément à d'autres travaux en langue arabe, la rime ne semble pas constituer une unité pertinente lors de la lecture (Ammar, 2003 ; Boukadida, 2008 ; Besse & al., 2007). Compte tenu de la situation de diglossie, le développement d'une capacité précoce d'analyse en attaque et rime est plus difficile en langue arabe. La structure morphologique particulière de la l'arabe fondée sur la racine consonantique, invariable lors du passage de l'arabe dialectal à l'arabe standard, laisse supposer que les mots sont plus facilement décomposables en racine consonantique et morphèmes vocaliques qu'en attaque et rime (Ammar, 1997). Muter & al. (2004) ont montré, de leur côté, que les capacités de repérage phonémique et de repérage de rimes sont clairement dissociables. Leurs résultats confirment que seules les performances à des tâches métaphonémiques sont très fortement corrélées au niveau de lecture en 1ère année de scolarité contrairement aux performances à des tâches de manipulation de rimes.

Le niveau d'identification de mots à T2 est associé également aux épreuves de connaissance des lettres, de compréhension de consignes et aux connaissances morpho-syntaxiques. La connaissance des lettres est un facteur important dans le développement des procédures d'identification de mots. En favorisant l'apprentissage du son des lettres et le développement de la sensibilité phonémique, cette connaissance contribue à l'accès au principe alphabétique et à la maîtrise des habiletés de décodage phonologique, deux composantes centrales de la réussite dans l'apprentissage de la lecture (Foulin, 2007 ; 2005 ; Foulin & Pacton, 2006). La compréhension de consignes semble également très liée à la réussite à l'épreuve de lecture à T2. En effet, c'est sur la base de sa compréhension de la consigne que l'enfant va devoir réaliser la tâche demandée. Ainsi, des difficultés repérées dans certaines épreuves peuvent être dues à des difficultés liées à la compréhension des consignes qui les introduisent. Cette compétence mérite donc d'être étudiée davantage en CP puisque la majorité des activités

scolaires est introduite par des consignes simples et complexes. En 1ère année, la tâche de jugement des liens morphologiques présente le coefficient de corrélation le plus élevé avec les scores à l'épreuve IME1 ($r=.46$; $p<.01$).

L'analyse de régression sur l'épreuve d'identification de mots en 1ère année confirme la forte valeur prédictive de la connaissance précoce des lettres et de la suppression syllabique. Ceci est en accord avec les résultats de nombreux travaux qui ont montré le rôle important des facteurs langagiers, en particulier la conscience phonologique et la connaissance des lettres dans la prédiction des performances ultérieures en lecture (Biot-chevrier & al., 2008 ; Lecoq, 1991 ; Muter & al., 2004 ; Schatschneider & al., 2004).

La présente étude longitudinale montre également le poids important des compétences morpho-syntaxiques en CP qui prédisent une part propre de la variance en lecture, une année plus tard. Ce résultat met en évidence la contribution des connaissances morphologiques flexionnelles à l'apprentissage de la lecture en langue arabe. Les études qui ont porté sur les liens entre conscience morphologique et apprentissage de la lecture se sont focalisées, pour une grande part, sur le domaine de la morphologie dérivationnelle (Boukadida, 2008 ; Carlisle, 1995 ; Marec-Breton, 2003). Les recherches concernant la morphologie flexionnelle ont étudié surtout les liens entre les compétences morpho-syntaxiques et l'installation des habiletés orthographiques (Fayol & Jaffré, 2008). Les caractéristiques orthographiques de l'arabe font que l'analyse morphologique est nécessaire lors de la reconnaissance des mots écrits, et qu'elle contribue à l'explication de la variance des performances en lecture.

Prédicteurs de l'apprentissage de la lecture en 2ème année (T2)

Les analyses de corrélation entre les mesures en CP (T1) et le niveau d'identification de mots en 2ème année (T3) montrent une certaine stabilité. En fonction du niveau scolaire, certaines corrélations se renforcent légèrement, notamment en ce qui concerne certaines tâches phonologiques (comptage et segmentation syllabiques, détection de l'intrus syllabique) alors que d'autres s'affaiblissent telles que la suppression syllabique. Ceci est le résultat semble-t-il de l'influence de l'apprentissage de la lecture et des exercices scolaires systématiques. Etant donné les progrès réalisés par les élèves dans le traitement de l'information orale et écrite, certaines épreuves perdent de leur valeur discriminative, ce qui pourrait expliquer cette baisse légère du nombre des corrélations significatives.

Certaines mesures apparaissent corrélées davantage avec les scores finaux au Timé2 qu'avec ceux obtenus à l'épreuve IME2, c'est le cas notamment de la compréhension de consignes et de la suppression syllabique. En ce qui concerne l'épreuve de compréhension, les meilleures corrélations sont obtenues avec les épreuves de compréhension de consignes, de morpho-syntaxe et de détection de l'intrus syllabique. Contrairement à nos attentes, la compréhension orale en CP ne semble pas fortement corrélée à la compréhension écrite en 2ème année[71]. Cette faible relation entre les deux compétences peut s'expliquer par la différence entre la langue orale (arabe dialectal) utilisée par les élèves en CP et la langue écrite (arabe standard) qu'ils vont découvrir dans leur livre de lecture. En effet, l'épreuve de compréhension orale est passée en arabe dialectal alors que l'épreuve de compréhension écrite est composée de phrases en arabe standard. Il existe ainsi un écart important entre les constructions syntaxiques utilisées à l'oral par les élèves et celles qui organisent les premières phrases soumises à leur lecture.

Nous avons constaté également une corrélation significative entre les connaissances morphologiques mesurées à T2 et le niveau de lecture à T3 (IME2 : r=.37 ; Timé2 : r=.46 ; Cph : r=.34 ; p<.05). Les performances des élèves dans la tâche de morphologie dérivationnelle en 1ère année sont même plus corrélées avec les performances en lecture une année plus tard que ne le sont les performances dans la manipulation des unités phonémiques. Ces résultats sont en accord avec ceux trouvés par Boukadida (2008).

L'analyse de régression sur les épreuves d'identification de mots en 2ème année (IME2 et Timé2) montre la stabilité prédictive de la connaissance des lettres et des connaissances phonologiques. En revanche, c'est la segmentation syllabique qui apporte une contribution significative à l'explication de la variance en identification de mots en 2ème année. A ce niveau, les connaissances morpho-syntaxiques ne s'avèrent pas liées au niveau d'identification de mots mais s'investissent uniquement au niveau des capacités de compréhension. Le niveau de compréhension en 2ème année est expliqué en premier lieu par les connaissances morpho-syntaxiques (15% de la variance expliquée). En effet, la compréhension d'une phrase exige que les mots qui la constituent soient identifiés mais aussi que les relations entre les mots soient saisies. Des connaissances morpho-syntaxiques doivent donc intervenir dans la lecture-compréhension. Ces données confirment les résultats des recherches qui ont mis en évidence une relation étroite entre compréhension et niveau de

[71] Nous avons utilisé la même épreuve : en version orale en CP (arabe dialectal) et en version écrite en 2ème année (arabe standard).

conscience morphologique (Lecoq & al., 1996). En effet, la conscience morphosyntaxique rend compte de la manipulation d'éléments véhiculant des informations syntaxiques et sémantiques impliquées de façon cruciale dans le traitement des phrases (Colé & Fayol, 2000).

L'ensemble de ces résultats suggère que la connaissance des lettres semble jouer un rôle déterminant pour le développement des processus d'identification de mots. De même, Muter & al. (1998) montrent qu'une part de la variance des performances en lecture en 1ère année est expliquée par la connaissance des lettres, puis par la conscience phonémique. En 2ème année, les performances en lecture sont expliquées par celles obtenues l'année précédente et également par la connaissance des lettres au même moment.

Pour ce qui concerne l'ensemble des épreuves phonologiques, les résultats montrent une fois de plus le poids de ces compétences dans la prédiction des scores ultérieurs en lecture et reproduisent largement ceux observés par d'autres auteurs qui ont entrepris une étude de même type (Muter & al., 1998 ; 2004 ; Lecoq, 1991 ; Piquard-Kipffer, 2003 ; Shatschneider & al., 2004). Nous constatons en particulier que les résultats varient en fonction des unités considérées dans ces tâches (syllabes, rimes ou phonèmes) et également en fonction du type d'opérations effectuées (comptage, segmentation, suppression, détection de l'intrus, etc.). Nous remarquons que la modalité de passation (individuelle ou collective) intervient également puisque ce sont les épreuves individuelles exigeant une manipulation des unités syllabiques (comptage, segmentation et suppression syllabique) qui sont extraites dans les analyses de régression. La contribution des connaissances phonologiques à l'explication de la variance des capacités de compréhension apparaît moins importante que leur contribution aux capacités d'identification de mots. C'est l'épreuve de détection syllabique qui ajoute une part explicative non négligeable à la variance des performances en lecture-compréhension (10% de variance expliquée).

Par ailleurs, nous observons très peu de liaisons entre la vitesse de dénomination et les différentes épreuves de lecture. Cette compétence n'apporte aucune contribution significative à l'explication des performances lexiques. Ceci est en contradiction avec les résultats présentés par d'autres auteurs qui mettaient en évidence le rôle important de la dénomination rapide ou RAN (Rapid Automatized Naming) dans la prédiction du niveau ultérieur en lecture (Castel & al., 2008 ; Kirby & al., 2003 ; Lecoq, 1991 ; Scarborough, 1998). L'absence de relation entre les capacités d'accès au lexique et les performances en lecture peut être

expliquée par les caractéristiques de la langue arabe et plus particulièrement du contexte d'apprentissage. Compte tenu de la situation de diglossie, l'enfant tunisien aurait des difficultés à accéder de façon rapide et précise aux représentations phonologiques en arabe standard. En effet, lorsqu'il aborde la première année d'apprentissage, il dispose de représentations phonologiques de sa langue maternelle, l'arabe dialectal. Cependant, lors de la passation de cette épreuve, les enfants avaient très souvent tendance à utiliser le registre de l'arabe standard qui est associé à l'école et aux apprentissages scolaires. Ils passent ainsi plus de temps pour trouver le mot dans leur lexique mental. Certains auteurs suggèrent en effet que les processus mobilisés par ces épreuves interviendront à un niveau plus élevé et que la valeur prédictive de cette variable augmente avec l'avancement dans la scolarité (Kirby, Parilla & Pfeiffer, 2004).

Parmi les variables mesurées à T2, c'est le jugement des liens morphologiques qui apporte la contribution la plus importante à l'explication de la variance des capacités d'identification de mots et de compréhension (IME2 : 13% ; Timé2 : 21% ; compréhension : 11% de variance expliquée). Cette intervention est plus importante que celle des connaissances phonémiques. Ce résultat concorde avec ceux obtenus en français par Casalis & Louis-Alexandre (2000) qui montrent que les connaissances morphologiques affectent les performances en lecture de mots et en compréhension. Au niveau du cours préparatoire, aucune des tâches métamorphologiques utilisées ne rend compte des performances à la tâche de lecture de mots, seule la tâche de suppression de phonème est liée à la lecture. Par contre, au CE1, les réussites dans la tâche de production de mots dérivés et fléchis sont significativement liées aux performances de la lecture de mots et celles dans une tâche de segmentation morphologique à la compréhension en lecture. Ce résultat vient confirmer également celui obtenu par Boukadida (2008) en arabe : les connaissances morphologiques dérivationnelles précoces sont plus corrélées avec les performances en lecture une année plus tard que ne le sont les performances dans la manipulation des unités sonores.

En début d'apprentissage de la lecture, la morphologie contribue, à elle seule, à l'explication de la variance en lecture. La détection de l'intrus de phonème initial contribue également à expliquer les capacités d'identification de mots (ajoute 4% de variance à l'épreuve IME2 et 7% au Timé2) et de compréhension (2% de variance expliquée). La détection du phonème final contribue seulement à expliquer les capacités de compréhension (6% de variance expliquée).

Les données recueillies permettent de confirmer notre première hypothèse postulant un rôle important des compétences langagières dans la prédiction des performances en lecture et le poids plus modeste des facteurs cognitifs. Bien que ce ne soit pas toujours les mêmes variables qui sont extraites en fonction d'une part du niveau scolaire et d'autre part de l'épreuve de lecture utilisée, nous avons observé en particulier que les épreuves langagières ont un poids plus important dans la prédiction des performances ultérieures en lecture. C'est l'épreuve de connaissance des lettres qui est la plus fréquemment extraite dans les analyses en première position, quelle que soit l'épreuve de lecture (IME1, IME2 et Timé2). Les meilleures prédictions sont obtenues avec le Timé2 (26% de variance expliquée à cette épreuve). Ainsi, nous pouvons dire que la contribution de cette variable se maintient dans le temps.

Il semble, par conséquent, que les variables cognitives aient un poids moins déterminant lors des premiers apprentissages de la lecture. Nous retrouverons là également les principaux résultats obtenus par différents auteurs qui ont montré de faibles corrélations entre les capacités cognitives, plus particulièrement le niveau intellectuel, et les performances observées lors des premières étapes de l'apprentissage de la lecture, même si cette liaison devient de plus en plus forte au fur et à mesure qu'on avance dans le cursus scolaire (Lecoq, 1991 ; Siegel, 1990 ; 1989 ; 1988 ; Jimenez & al., 2003). Nous pouvons alors supposer que les contraintes scolaires et l'accroissement de la compétence lexique développeraient les compétences cognitives des enfants qui prendraient alors, à des étapes ultérieures, une place plus importante dans l'explication des différences individuelles en lecture. Il sera alors intéressant de suivre cette évolution et de voir quelle sera l'orientation que prendront les corrélations entre les différentes épreuves.

Les données de cette recherche montrent une certaine variabilité de la prédiction en fonction de l'épreuve de lecture utilisée en tant que variable dépendante. Les meilleures prédictions du niveau en identification de mots en 2ème année sont obtenues avec le Timé2 (37% de variance expliquée *vs* 27% pour IME2). Les deux épreuves d'identification de mots écrits utilisées dans cette recherche présentent des ressemblances : les deux sont basées sur des scores composites et sont administrées en passation collective. Toutefois, la nature des compétences évaluées en lecture diffère d'une épreuve à l'autre, ce qui peut expliquer cette variation. A l'épreuve IME, la tâche de l'enfant consiste à vérifier chaque fois l'adéquation entre une image et un mot qui peut être soit correct soit incorrect. Ce sont les items comportant une perturbation au niveau des voyelles (items PLV) qui s'avèrent difficiles pour les enfants

(tableau 9 et figure 7). Compte tenu de la préeminence des consonnes par rapport aux voyelles dans la langue arabe, l'enfant ne se rend pas compte des erreurs portant sur les signes vocaliques. Dans les trois tâches du Timé2, l'enfant doit chercher chaque fois le mot cible parmi une liste de 5 items tests. Ainsi, la tâche de l'enfant est facilitée par la présence du mot correct. De plus, dans le Timé2.3, la tâche de l'enfant dépasse le simple décodage et nécessite une capacité de catégorisation sémantique. L'analyse de la variance et de la distribution des scores obtenus à ces deux épreuves montre également des différences : une distribution plutôt symétrique (courbe normale) pour l'épreuve IME2 et une distribution asymétrique (courbe en J) pour le Timé2 (voir annexe III). Contrairement à l'épreuve IME2, l'épreuve Timé2 n'a pas fait l'objet d'un étalonnage sur la population tunisienne. Les résultats obtenus pourraient ainsi contribuer à la construction ultérieure d'un instrument plus élaboré.

L'ensemble de ces résultats pourrait suggérer de développer un modèle des prédicteurs de l'apprentissage de la lecture en langue arabe. Le chapitre suivant a pour objectif de proposer un modèle structurel des relations prédictives entre les connaissances phonologiques et morphologiques précoces et les performances en lecture.

Chapitre 3 : Modèle d'équations structurelles des prédicteurs de l'apprentissage de la lecture en arabe

Ce chapitre a pour objectif de proposer un modèle des prédicteurs de l'apprentissage de la lecture en langue arabe. Un modèle d'équations structurelles (MES) a pour but de traiter statistiquement des relations de causalité hypothétiques entre plusieurs variables. Ces modèles apportent une distinction entre variables latentes et variables observées[72]. Les variables latentes renvoient à des construits psychologiques non observables directement. Elles sont appréhendées par un ensemble d'indicateurs constituant les variables observées.

Les MES peuvent être considérés comme l'association de deux modèles distincts : un modèle de mesure qui représente la structure des liaisons entre les variables latentes et leurs indicateurs observés et un modèle de relations linéaires structurelles qui établit les liaisons supposées causales entre, d'une part, les variables latentes explicatives et les variables latentes à expliquer et d'autre part entre les différentes variables latentes à expliquer. Le modèle de mesure précise les indicateurs (variables observées) de chaque variable latente. Chaque indicateur peut être défini en reprenant la structure factorielle apparue à l'issue d'une analyse en composantes principales. Le modèle structurel concerne les hypothèses de relations linéaires entre les variables latentes.

Notre objectif est d'établir un modèle des relations longitudinales entre les différentes mesures à T1 et T2 et les performances en lecture à T3 (identification de mots et compréhension). Dans cette perspective, nous allons tester un modèle intégrant un ensemble de variables indépendantes susceptibles de prédire les performances en lecture en suivant la démarche méthodologique établie par Roussel & al. (2002). D'abord, une analyse factorielle confirmatoire a été réalisée afin de vérifier la validité des construits et ainsi confirmer les facteurs recherchés. La deuxième étape consiste en une analyse structurelle qui permettra d'évaluer dans quelles mesures le modèle théorique reproduit correctement les données collectées (*cf* . annexe IV pour une synthèse sur la démarche méthodologique des MES).

[72] L'intérêt de cette distinction entre variables latentes et observées est que l'erreur de mesure va pouvoir être évaluée dans le modèle.

1. Spécification du modèle théorique

La problématique de cette recherche porte sur la prédiction de l'apprentissage de la lecture en langue arabe. Nous proposons une première modélisation des relations prédictives entre les variables mesurées à T1 et T2 et les performances en lecture à T3 en s'appuyant sur divers travaux. La première base théorique est celle des résultats des études prédictives récentes (*cf.* chapitre 4) qui ont essayé de déterminer les compétences précoces susceptibles de prédire le niveau ultérieur en lecture. D'après ces études, les compétences en langage oral sont les facteurs les plus prédictifs du niveau ultérieur de lecture. Le langage oral est conçu comme étant la résultatnte de l'intégration de plusieurs sous-composantes (phonologie, morphologie, syntaxe, etc.). Ces recherches confirment ainsi l'importance des compétences phonologiques et morphologiques, de la connaissance des lettres et de la dénomination rapide. Les travaux interlangues (Besse, Demont & Gombert, 2007 ; Georgiou, Parilla, Papadopoulos, 2008) et les études sur la langue arabe (Abu-Rabia, 2007 ; Ammar, 2003 ; Boukadida, 2008 ; Boukadida & al., 2009 ; Besse, 2007) constituent une seconde base théorique. Ces études confirment le rôle important des connaissances phonologiques dans l'apprentissage de la lecture en arabe mais révèlent également des différences liées aux caractéristiques de la langue arabe et du contexte d'apprentissage. Elles ont démontré ainsi le rôle important des connaissances morphologiques dans cet apprentissage.

De nombreux travaux de recherche réalisés dans ce domaine ont permis ainsi de préciser les compétences langagières et cognitives impliquées dans l'acquisition de la lecture et de développer des modèles explicitant les relations entre ces différentes dimensions (Ecalle & Magnan, 2002 ; Sprenger-Charolles & Colé, 2003). Storch & Witehurst (2002) ont essayé de modéliser les liens prédictifs entre les compétences langagières orales et les capacités ultérieures en lecture en adoptant une approche structurelle. Les résultats du MES montrent que les habiletés langagières orales précoces n'apparaissent pas avoir une contribution directe à la capacité ultérieure de lecture au cours des deux premières années de l'enseignement formel. Toutefois, cette relation s'exprime de façon indirecte par le biais des habiletés grapho-phonologiques considérées comme une variable médiatrice. Les résultats obtenus apportent un support empirique à l'idée selon laquelle les compétences langagières (connaissances des lettres, niveau du vocabulaire) et les compétences de recodage phonologique jouent un rôle significatif à différents moments du développement de l'acquisition de la lecture. Dans la même perspective, Muter & al. (2004) ont étudié la relation prédictive entre les compétences

langagières précoces (compétences phonologiques, connaissance des lettres, connaissances grammaticales et niveau du vocabulaire) et les capacités d'identification de mots et de compréhension au cours des deux premières années de l'apprentissage de la lecture. Les résultats du modèle en pistes causales[73] montrent que les capacités d'identification de mots sont prédites par les processus phonologiques, plus particulièrement la sensibilité phonémique et la connaissance des lettres qui expliquent une part importante de la variance, alors que la compréhension apparaît dépendante du niveau du vocabulaire et des connaissances morphologiques. Leong, Hau, Cheng & Tan (2005) ont essayé d'examiner l'effet de trois construits ou variables latentes (connaissances orthographiques, sensibilité phonologique et identification de mots) au temps 1 (T1) sur les mêmes performances au temps 2 (T2) chez 108 élèves chinois âgés de 10-11 ans. Les résultats de l'analyse en équations structurelles indiquent que les capacités d'identification de mots à T2 sont mieux prédites par le même construit à T1 alors que les connaissances orthographiques et la sensibilité phonologique ne semblent pas apporter une contribution significative à la variance.

En nous appuyant sur ces travaux, nous proposons une première modélisation des relations prédictives entre les différentes variables étudiées et l'efficience ultérieure en lecture. Le modèle proposé considère que les performances en lecture en langue arabe sont expliquées par deux variables latentes : connaissances phonologiques et connaissances morphologiques. Les connaissances phonémiques à T2 sont considérées comme une variable médiatrice qui prédit à son tour le niveau de lecture à T3. Le poids prédictif des connaissances phonologiques s'exprimerait ainsi directement sur les performances en lecture et de façon indirecte par le biais des connaissances phonémiques. Une variable médiatrice permet d'expliquer comment s'opère la relation entre une variable indépendante et une variable dépendante en décomposant cette relation en effet direct et en effet indirect médiatisé (Roussel & Wacheux, 2006). Ainsi, une première variable indépendante (X) influe sur une seconde variable intermédiaire (M), qui influe à son tour sur une variable dépendante (Y). L'effet indirect est évalué à partir du produit des coefficients de pistes entre XM et XY.

Nous proposons donc d'intégrer la variable « connaissances morphologiques » au modèle des prédicteurs de l'apprentissage de la lecture de façon à déterminer dans quelle mesure elle est liée aux performances en lecture et si elle joue un rôle aussi important que celui des

[73] Il y a une différence entre un modèle d'équation structurelle (MES) et un modèle en piste causale. Le premier traite les relations de causalité entre des variables latentes et observées alors que le second étudie seulement les relations causales entre variables observées.

connaissances phonologiques. Nous considérons également que les connaissances morphologiques seraient liées aux connaissances phonologiques. Cette décision est basée sur les résultats de plusieurs recherches qui ont mis en évidence une relation étroite entre ces deux compétences (Boukadida, 2008 ; Carlisle & Nomanbhoy, 1993 ; Casalis & Colé, 2009 ; Marec-Breton, 2003). L'ensemble de ces relations sera testé grâce à un modèle en équations structurelles, lequel permettra de répondre à ces différents objectifs. La figure ci-après résume les relations hypothétiques entre l'ensemble de ces variables.

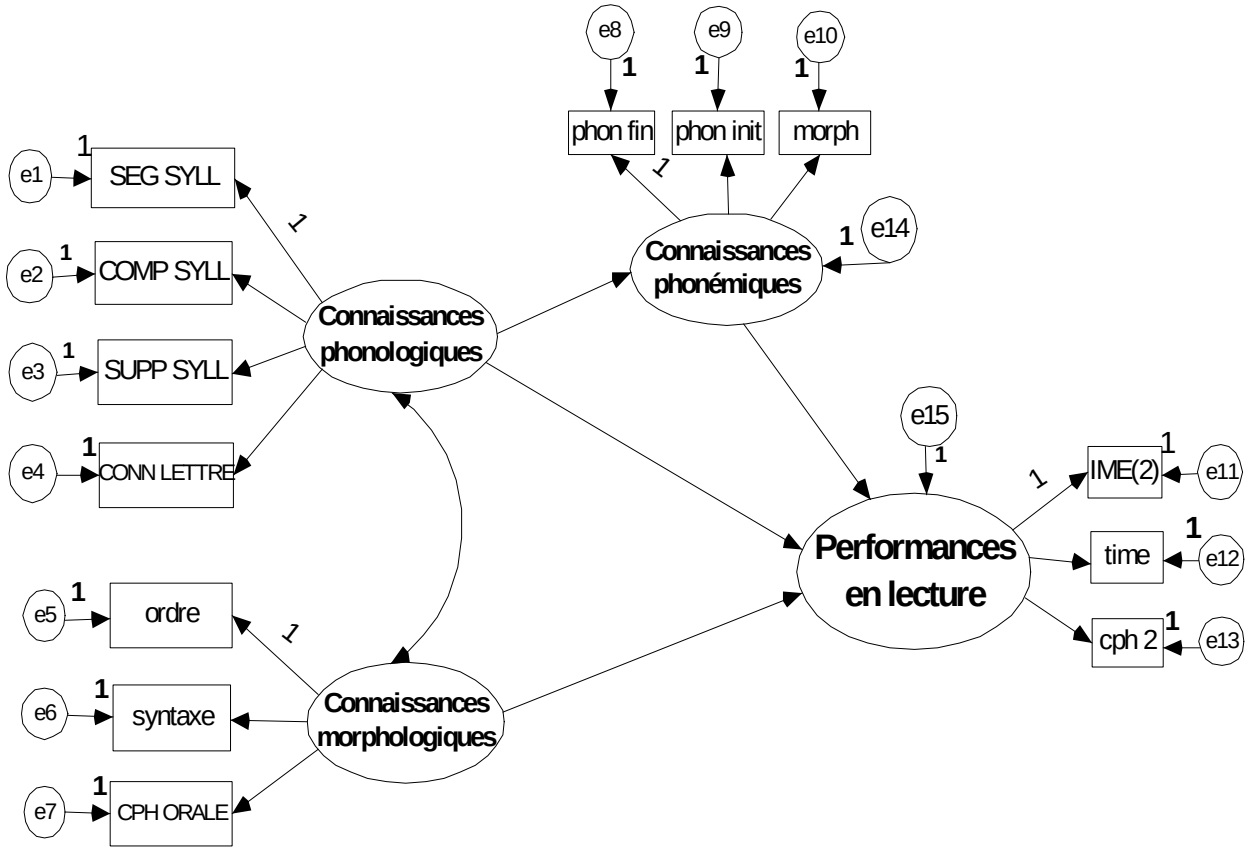


Figure 10 : Modèle théorique des prédicteurs de l'apprentissage de la lecture en langue arabe

Nous allons procéder en deux étapes. La première consiste en une estimation du modèle de mesure par une analyse factorielle confirmatoire (AFC) afin d'évaluer la validité convergente des construits. La seconde est consacrée à l'estimation et au test des relations structurelles entre les différentes variables latentes.

2. Estimation du modèle de mesure

Le modèle de mesure décrit les relations entre les variables latentes et les variables observables ou mesurées. Pour pouvoir spécifier les relations entre les concepts latents et les variables observées, nous avons construit un modèle de mesure composé de quatre variables latentes : les connaissances phonologiques, les connaissances morphologiques, les connaissances phonémiques et les performances en lecture. Il est fortement recommandé d'avoir au minimum trois indicateurs par variable latente. L'utilisation d'un indicateur unique entraîne toujours un problème d'identification (Roussel & al., 2002). Avec une seule variable observée par facteur, l'utilisateur est incapable d'estimer l'erreur aléatoire dans le modèle de mesure. Cette contrainte nous a empêchés d'introduire certaines variables latentes dans le modèle pour lesquelles nous avons un seul indicateur tel que le niveau d'identification de mots en 1ère année.

Les indicateurs des variables latentes sont ceux dégagés à partir de l'analyse en composantes principales réalisée dans le chapitre précédent. Ainsi, les indicateurs de la variable latente « connaissances phonologiques » sont les scores moyens obtenus aux épreuves de comptage, segmentation et suppression syllabiques et à l'épreuve de connaissance des lettres. Les indicateurs du construit ou variable latente « connaissances morphologiques » sont représentés par les scores moyens obtenus aux tâches morpho-syntaxiques (jugement de l'ordre et morpho-syntaxe) et à l'épreuve de compréhension orale. Les indicateurs du construit « connaissances phonémiques » sont les scores obtenus aux épreuves phonémiques et à l'épreuve de jugement des liens morphologiques à T2. Les indicateurs de la variable latente à expliquer « performances en lecture à T3 » correspondent aux scores moyens obtenus par les enfants aux épreuves d'identification de mots (IME2 et Timé2) et à l'épreuve de compréhension. Chaque variable observée est donc expliquée par une seule variable latente et une erreur de mesure qui lui est propre. Afin d'éviter les problèmes de sous-identification, il est souhaitable pour la mesure de chaque variable latente de fixer la contribution factorielle de son premier indicateur à 1. Dans le but de valider ce modèle de mesure, nous allons examiner sa validité et sa qualité d'adéquation aux données à l'aide de l'analyse factorielle confirmatoire (AFC). L'analyse factorielle confirmatoire, appliquée au modèle de mesure, consiste à spécifier les relations entre les variables latentes et les indicateurs observés. Cette analyse est destinée à vérifier que les construits sont clairement identifiés par les facteurs communs (Roussel & Wacheux, 2006). Trois étapes sont nécessaires. La première est

l'examen de la signification statistique des contributions factorielles des indicateurs, la seconde est l'évaluation de la variance expliquée et enfin la vérification de l'adéquation du modèle théorique aux données empiriques.

2.1. Estimation de la validité

L'AFC a pour objet de tester la validité d'un modèle de mesure, en vérifiant que le modèle postulé correspond aux données observées. Roussel & al. (2002) recommandent d'utiliser l'AFC afin de procéder à des tests de validité de construit permettant d'éprouver la validité convergente d'une échelle de mesure qui se vérifie à partir des contributions factorielles. La validité convergente peut être vérifiée si et seulement si le test t (C.R) est supérieur à 1,96 au niveau de signification de 5% pour chaque contribution factorielle vérifiant ainsi la signification du lien de chaque indicateur à son construit.

Tableau 36 : Estimation de la validité convergente du modèle de mesure

Poids factoriels standardisés	**Estimation**	**C.R.**
Seg syll <--- Conn phono	.981	11.825*
Compt syll <--- Conn phono	.949	12.072*
Supp syll <--- Conn phono	.567	8.657*
Conn let <--- Conn phono	.452	6.448*
Jug ordre <--- Conn morpho**	.635	-
Morpho-Syn <--- Conn morpho	.707	5.837*
Cph orale <--- Conn morpho	.410	4.299*
Phonè init <--- Conn phonémiques	.777	4.938*
Phonè fin <--- Conn phonémiques**	.628	-
IME2 <--- performances lecture**	.793	-
Timé2 <--- performances lecture	.961	12.542*
Cph <--- performances lecture	.698	9.786*

*p<.05

** la valeur du C.R de ces paramètres est impossible à calculer puisqu'ils sont fixés à 1.

Il ressort d'après les résulatas présentés dans le tableau 36 que chaque construit montre une bonne validité convergente. Nous pouvons dire que les contributions factorielles des indicateurs sont globalement satisfaisantes pour tous les construits. En effet, les coefficients de saturation entre les indicateurs et les variables latentes se sont révélés positifs. Le test t

associé à chacune des contributions factorielles est significatif pour toutes les épreuves, condition nécessaire à la validité convergente (supérieur à 1,96).

Cette condition est vérifiée pour le construit « connaissances phonologiques ». Ce sont les épreuves de comptage et de segmentation syllabique qui apportent les meilleures contributions factorielles (respectivement λ =.949 et λ =.981 ; au seuil de 5%). Concernant le construit « connaissances morphologiques », nous remarquons que les trois indicateurs apportent des contributions factorielles significatives : jugement de l'ordre (λ =.635 ; au seuil de 5%), la morpho-syntaxe (λ =.707 au seuil de 5%) et la compréhension orale (λ =.410 ; au seuil de 5%). L'analyse des contributions factorielles relatives au construit « performances en lecture à T2 » montre par ailleurs qu'il est bien représenté par les différents indicateurs : identification de mots à T2 (λ=.793 ; au seuil de 5%), les performances au Timé2 (λ =.961 ; au seuil de 5%) et la compréhension écrite (λ =.697 ; au seuil de 5%).

Un autre critère de qualité d'ajustement examiné est celui des corrélations entre les variables latentes explicatives. L'analyse des corrélations entre les connaissances phonologiques et morphologiques montre une relation significative (tableau 37) mettant en évidence un chevauchement de ces deux construits.

Tableau 37 : Corrélation entre les variables latentes

	Estimation
Connaissances phono – Connaissances morpho	.409

Ces résultats confirment bien que le modèle de mesure est acceptable et permettent de conclure que les variables exogènes ou construits sont relativement bien expliquées par les variables endogènes (indicateurs) qui leurs sont associées.

2.2. Evaluation de la variance expliquée

Le pourcentage de variance expliquée ou R^2 (Squared Mutiple corrélations)[74] s'interprète comme le pourcentage de variance de la variable à expliquer restituée par le modèle (tableau 38). Plus la somme des R^2 est élevée, plus la variance des variables expliquées restituée par le modèle est élevée. L'examen des pourcentages de variance (R^2) relatives au premier construit

[74] Certains auteurs estiment que des coefficients R^2 situés aux alentours de .20 sont acceptables, étant données la complexité des phénomènes étudiés en sciences sociales et la multiplicité des facteurs qui influencent ces phénomènes (Roussel & al., 2002).

« connaissances phonologiques » sont respectivement de l'ordre de 97% pour la segmentation syllabique, 89% pour le comptage et 32% pour la suppresion. Ces résultats attestent d'un important pouvoir explicatif. Nous relevons également que les parts de variance expliquées par les indicateurs du construit « connaissances morphologiques » sont satisfaisantes. En effet, le jugement de l'ordre restitue 51% de la variance alors que la morpho-syntaxe explique jusqu'à 54% de la variance. Le pourcentage de variance expliquée de la compréhension orale est faible (17%).

Tableau 38: Estimation de la variance expliquée

R^2	**Estimation**
Seg syll	.975
Compt syll	.898
Supp syll	.325
Conn let	.207
Jug ordre	.519
Morpho-Syn	.540
Cph orale	.174
Phonè init	.456
Phonè fin	.324
Jug morph	.245
IME2	.616
Timé2	.923
Cph	.472

Concernant la variable latente « connaissances phonémiques », le pourcentage de variance expliquée est de 32% pour la détection du phonème final, 45% pour la détection du phonème initial et 24% pour le jugement des liens morphologiques.

Enfin, pour la variable « performances en lecture », on constate que les variances expliquées des trois indicateurs sont importantes et significatives. En effet, la part de variance expliquée de l'identification de mots est de 61%, celle du Timé2 est de 92% et celle exprimée par rapport à la compréhension est de 47%. Nous jugeons ces résultats satisfaisants. Ces données indiquent que les variables exogènes (construits) sont relativement bien expliquées par les variables endogènes (indicateurs) qui leurs sont associées.

2.3. Evaluation de la qualité d'ajustement

Après avoir estimé les paramètres du modèle, l'évaluation des critères du bon ajustement est la phase essentielle afin de déterminer dans quelle mesure les données recueillies sont en adéquation avec le modèle postulé. La méthode d'estimation qui est utilisée pour calculer les équations structurelles est celle du maximum de vraisemblance (Maximum Likelihood) qui tolère des violations modérées de la normalité et fournit de meilleurs indices d'ajustement.

Le niveau d'ajustement du modèle est évalué par la statistique Chi-deux ($\chi 2$). L'ajustement du modèle aux données est considéré comme adéquat lorsque la valeur du Chi-deux normé ($\chi 2$/dl) est inférieure à 5 (Roussel & & al., 2002). Le $\chi 2$ est le plus souvent complété par divers indices d'ajustement qui sont plus pratiques et plus robustes pour indiquer à quel point le modèle explique les données. Dans cette perspective, nous avons retenu les indices les plus couramment utilisées dans la recherche : les indices absolus[75] (χ^2, GFI[76], AGFI[77] et RMSEA[78]), les indices incrémentaux[79] (NFI[80], CFI[81], IFI[82] et NNFI[83]), et les indices de parcimonie[84] ($\chi 2$/dl).

Le χ^2 représente un indicateur du niveau de correspondance entre une structure factorielle proposée (i.e., le modèle postulé) et les données collectées. Pour que le modèle soit accepté, le χ^2 doit être non significatif, ce qui signifie qu'il n'y a pas de différence entre le modèle postulé et les données. Le GFI indique la proportion de variance dont rend compte le modèle. Le NFI évalue l'ajustement du modèle proposé par rapport à un modèle théorique d'indépendance entre les variables. Une règle communément admise consiste à considérer le modèle comme correct quand les indices (compris entre 0 et 1) sont supérieurs à .90. Le RMSEA mesure l'erreur moyenne d'approximation attachée au modèle spécifié. Un RMSEA inférieur ou égal à.08 témoigne d'une qualité d'ajustement raisonnable. En revanche, tout modèle présentant un

[75]Un indice d'ajustement absolu permet d'évaluer dans quelle mesure le modèle théorique posée a priori reproduit correctement les données collectées.
[76] Goodness of Fit Index
[77] Adjusted Goodness of Fit Index
[78] Root Mean Square Error of Approximation
[79]Un indice incrémental mesure l'amélioration de l'ajustement en comparant le modèle testé à un modèle plus restrictif, dit « modèle de base ».
[80] Normed Fit Index
[81] Comparative Fit Index
[82] Incremental Fit Index
[83] Non-Normed Fit Index
[84]Les indices de parcimonie indiquent dans quelle mesure le modèle présente un bon ajustement pour chaque coefficient estimé.

RMSEA supérieur à.10 doit être rejeté. Le tableau 39 expose les indices d'adéquation pour le modèle présenté.

Tableau 39 : Indices d'ajustement du modèle de mesure

Indice	**χ^2**	**dl**	**χ^2/dl**	**GFI**	**AGFI**	**NFI**	**CFI**	**IFI**	**NNFI**	**RMSEA**
Valeur	288,3	64	4,505	.863	.833	.852	.882	.796	.846	,096
Seuil	ns	-	<5	>.90	>.90	>.90	>.90	>.90	>.90	<.08

Les indices obtenus ne sont pas satisfaisants si on les compare aux seuils proposés par Roussel & al. (2002). Nous décidons alors d'apporter des modifications au modèle pour améliorer la qualité d'ajustement. La première modification consiste à renoncer aux indicateurs qui présentent les paramètres (poids factoriel et variance expliquée) les plus faibles, telles que l'épreuve de compréhension orale et de connaissance des lettres, vu qu'elles présentent des erreurs de mesures importantes. Les nouveaux indices d'ajustement sont présentés dans le tableau 40.

Tableau 40 : Indices d'ajustement du modèle respécifié

Indice	**χ^2**	**dl**	**χ^2/dl**	**GFI**	**AGFI**	**NFI**	**CFI**	**IFI**	**NNFI**	**RMSEA**
Valeur	197,1	51	3,885	.899	.874	.867	.889	.841	.864	.078
Seuil	ns		<5	>.90	>.90	>.90	>.90	>.90	>.90	<.08

Les résultats montrent que les indices d'ajustement pour le modèle respécifié sont plus satisfaisants. Les valeurs du χ^2 et du RMSEA ont baissé. Cependant, les valeurs des autres indices (AGFI, NFI, CFI, IFI et TLI), qui doivent être idéalement supérieures à .90, sont moins satisfaisantes mais se rapprochent de ce seuil. Nous décidons de conserver ce modèle de mesure et de passer à l'estimation des relations structurelles entre les différentes variables latentes.

3. Estimation du modèle structurel

Une fois l'ajustement du modèle de mesure estimé, il est possible d'évaluer l'ajustement du modèle structurel. Il s'agit de vérifier que les relations entre les construits ou les variables latentes désignées par les coefficients structurels (appelés aussi coefficients de régression)

sont bien statistiquement significatives. Lorsqu'ils sont proches de 0, les effets sont faibles, tandis qu'en se rapprochant de 1, l'importance de la relation linéaire s'accroît.

Pour résumer, ce type d'analyse permet de distinguer le type d'effets (positifs ou négatifs) des variables entre elles, d'estimer le poids des variables indépendantes sur les variables dépendantes, et de déterminer par quelle voie (directe ou indirecte) chaque variable influence la suivante. Les tests des liaisons entre les variables du modèle structurel complet ont été effectués grace au logiciel AMOS. Les principaux résultats vont être présentés sous la forme d'un modèle schématisant les liens entre les variables. La figure ci-dessus résume l'ensemble des relations structurelles prises en compte dans le modèle et les coefficients de pistes standardisés associés à chacune de ces relations.

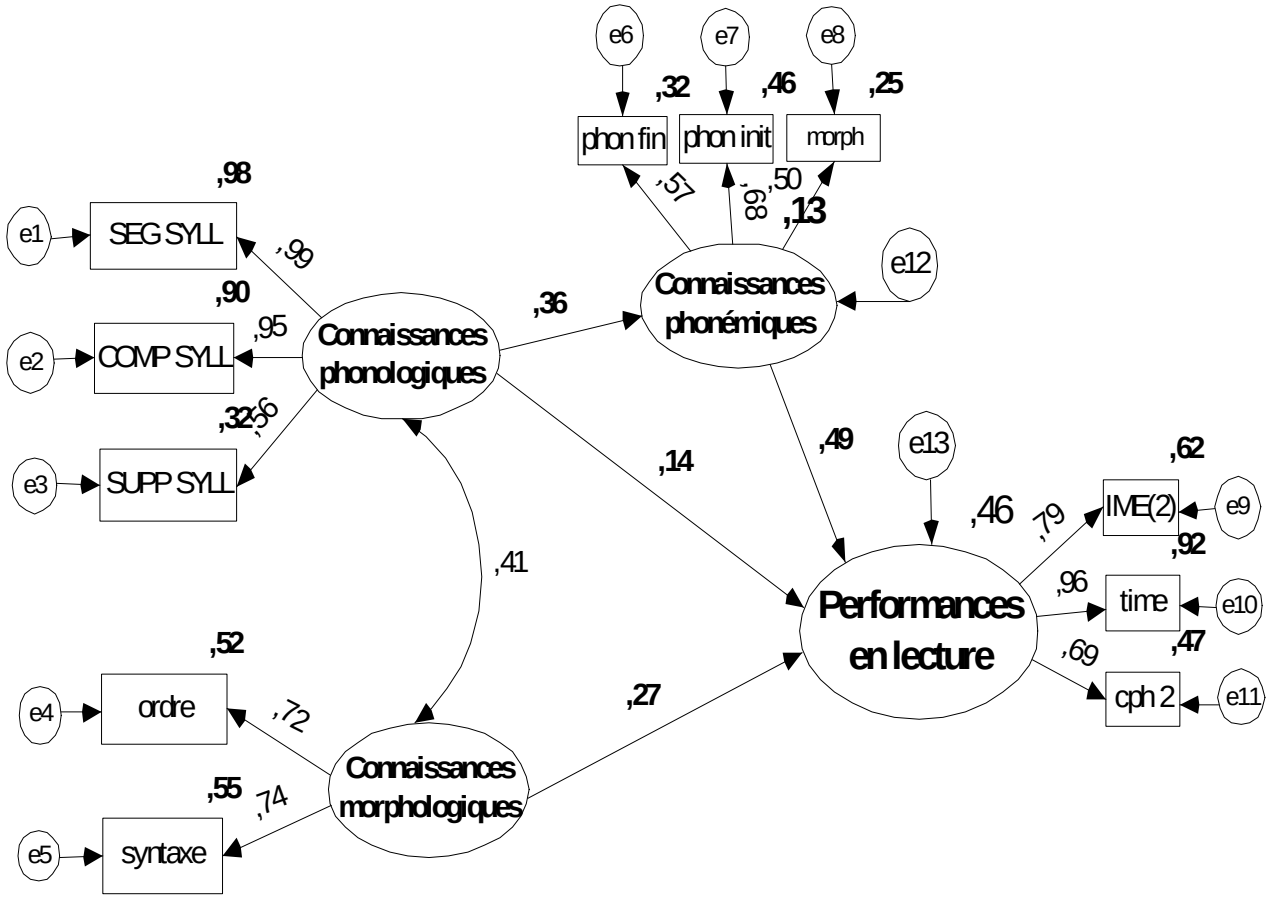


Figure 11 : Modèle d'équations structurelles des prédicteurs de l'apprentissage de la lecture en langue arabe

Dans le modèle d'équations structurelles, les construits ou variables latentes sont représentés par des formes ovales. Les rectangles représentent les indicateurs (variables observées)

utilisés pour mesurer les variables latentes. Les flèches signalent une relation de mesure entre les indicateurs et la variable latente. Les erreurs de mesure, dénommées « e_n », portent sur la part de variance d'une variable qui n'est pas expliquée par les autres variables du modèle. Toutes les pistes utilisées dans la construction du modèle sont représentées. Les chiffres au-dessus des flèches représentent les coefficients de pistes (β) standardisés et significatifs à $p < .05$ (voir figure 11). Les effets directs et indirects des variables latentes explicatives ou variables exogènes (connaissances phonologiques et morphologiques) sur la variable latente à expliquer ou variable endogène (performances en lecture) peuvent être évalués par les coefficients de causalité qui relient les différents construits de notre modèle à tester.

Le modèle d'équations structurelles présenté dans la figure 11 indique les relations prédictives entre les connaissances phonologiques et morphologiques à T1 et les performances en lecture à T3. Ce poids prédictif s'exprimerait directement, mais aussi de façon indirecte par le biais des connaissances phonémiques. Le niveau de conscience phonémique à T2 viendrait médiatiser le poids des connaissances phonologiques. Cette relation est définie comme une médiation ou bien un effet indirect (Preacher & Hayes, 2008).

3.1. Vérification de la significativité des paramètres du modèle

Afin de pouvoir confirmer que le modèle reproduit de manière correcte les données, il est nécessaire de vérifier la significativité des relations entre les variables latentes (Roussel et al., 2002). Seules les variables ayant une contribution significative (coefficient de régression entre les variables indépendantes et la variable dépendante), testées avec l'indice « t de student » (C.R.) sont conservées. Le tableau 41 présente un récapitulatif des résultats obtenus.

Tableau 41: Test de significativité des paramètres du modèle des prédicteurs de l'apprentissage de la lecture en arabe

Coefficients de régression standardisés	Estimation	C.R.
Lecture T2 <--- Conn phonémiques	.492	3.976*
Lecture T2 <--- Conn phonologiques	.143	1.634 ns
Lecture T2 <--- Conn morphologiques	.270	2.790*
Conn phonémiques <--- Conn phonologiques	.360	3.408*

*p<.05

L'analyse des coefficients de régression entre les différentes variables latentes explicatives et la performance en lecture montre que le coefficient de causalité reliant les connaissances

morphologiques à T1 et les performances en lecture à T3 est significatif (β=.27 ; t=2,790 au seuil de 5%). Les connaissances phonologiques ne paraissent pas avoir un effet direct sur la performance en lecture. En effet, la piste entre les deux compétences n'est pas statistiquement significative (β=.143ns). Par ailleurs, les connaissances phonémiques à T2 peuvent se conçevoir comme une variable médiatrice puisque les effets indirects des connaissances phonologiques sur les performances en lecture sont significatifs. Ainsi, les connaissances phonologiques n'affectent pas directement le niveau ultérieur de lecture, mais indirectement par le biais des connaissances phonémiques à T2. Afin de nous assurer de la significativité de l'effet médiateur, nous avons utilisé le test de Sobel (Preacher & Hayes, 2008) qui permet de calculer l'erreur standardisée de l'effet indirect[85].

Tableau 42: Estimation des effets directs et indirects des connaissances phonologiques sur les performances en lecture

Variables latentes	Effet direct	Effet indirect	Effet total
Connaissances phonologiques	.143ns	.176 Z= 5.80 ; p<.01	.319

Les résultats du test présentés dans le tableau 42 indiquent que les effets médiateurs sont statistiquement significatifs pour les liens entre les connaissances phonologiques et les performances en lecture. La médiation est partielle puisque la relation directe est non significative (effet direct = .143 ; effet indirect =.176). La somme des effets directs et indirects forme alors un effet total de .319. Ainsi, les connaissances phonologiques à T1 ont un pouvoir prédictif sur les performances en lecture à T3 qui s'exprime de façon indirecte par le biais des connaissances phonémiques à T2. A la lecture du tableau 41, nous constatons que les connaissances phonologiques à T1 exercent une influence sur les connaissances phonémiques à T2 à travers une relation significative positive (β=.360 ; t= 3,408 au seuil de 5%). De même, les résultats montrent qu'il existe un effet direct significatif entre les connaissances phonémiques et les performances en lecture (β= .492 ; t= 3,976 au seuil de 5%).

L'estimation des pourcentages de variance expliquée ou R^2 (Squared Multiple Correlations) montre que les variables du modèle expliquent 46% de la variance des performances en

[85] Ce test est disponible à l'adresse suivante : http://people.ku.edu/~preacher/sobel/sobel.htm

lecture. Les connaissances phonologiques expliquent, quant à elles, 13% de la variance des connaissances phonémiques.

3.2. Evaluation de la qualité d'ajustement

L'examen des différents indices d'adéquation (tableau 43) montre que le modèle est acceptable et qu'il s'ajuste suffisamment bien aux données empiriques pour que les estimations des paramètres soient fiables. La valeur de $\chi 2/dl$ est très satisfaisante, elle est inférieure à 5 (2,098). Les coefficients GFI et AGFI, très proches de .90, traduisent un très bon « fit » entre le modèle et les données (GFI=.935 et le AGFI=.878). Par ailleurs, l'indice RMSEA (en terme de variance résiduelle, c'est-à-dire de variance non expliquée), devant être idéalement inférieur à .05, est moins satisfaisant. Mais un RMSEA inférieur ou égal à 0.08 témoigne d'une qualité d'ajustement raisonnable. Les indices permettant de juger de la qualité d'ajustement du modèle, tels que le CFI et le NFI, ont des valeurs très proches de .90 (respectivement .965 et .936). De la même façon, l'IFI et le NNFI peuvent être qualifiés d'acceptables (respectivement .948 et .924).

Tableau 43 : Indices d'ajustement du modèle des prédicteurs de l'apprentissage de la lecture en arabe

Indice	χ^2	**dl**	$\chi 2/dl$	**GFI**	**AGFI**	**NFI**	**CFI**	**IFI**	**NNFI**	**RMSEA**
Valeur	88,685	38	2,334	.912	.874	.913	.947	.948	.924	,080
Seuil	ns	-	<5	>.90	>.90	>.90	>.90	>.90	>.90	<.05

4. Discussion

L'objectif principal de ce chapitre était d'apporter une première modélisation des prédicteurs de l'apprentissage de la lecture en langue arabe. Ainsi, nous avons essayé d'adopter une approche structurelle pour l'étude des liens prédictifs entre les compétences langagières mesurées à T1 et T2 et les performances en lecture à T3. Les MES sont élaborés pour tester les effets linéaires entre un ensemble de variables non directement observables : les variables latentes. Ces dernières sont estimées grâce à des variables manifestes ou observables : les indicateurs. Jusqu'à présent, les études portant sur l'apprentissage de la lecture s'étaient surtout intéressées à la dimension phonologique, au détriment d'autres dimensions. Cependant, les études interlangues concluent que les caractéristiques de la langue d'apprentissage constituent un facteur important devant être pris en compte dans les études

prédictives. Ainsi, il faut apprécier ce qui dépend des spécifités des langues ou des systèmes d'écriture.

Le recours à la méthode d'équations structurelles nous a permis de tester un modèle des prédicteurs de l'apprentissage de la lecture en langue arabe en nous référant à un ensemble de travaux de recherches qui ont essayé d'établir un lien entre certaines compétences langagières précoces et l'efficience ultérieure en lecture et de modéliser la relation prédictive entre les différentes dimensions (Storch & Witehurst, 2002 ; Muter & al., 2004 ; Leong & al., 2005). La majorité de ces recherches a insisté sur le rôle des compétences langagières orales, de la conscience phonologique et des connaissances orthographiques, mais aucune n'a pris en compte le rôle des connaissances morphologiques dans le modèle proposé. Les recherches réalisées sur la langue arabe ont confirmé le rôle primordial de la dimension morphologique dans l'apprentissage de la lecture. Les connaissances morphologiques expliquent une part importante de la variance en lecture en langue arabe (Abui-Rabia, 2007 ; Boukadida, 2008 ; Besse & al., 2007 ; Belajouza, 2003). Le modèle spécifié dans cette recherche est composé de quatre variables latentes : connaissances phonologiques, connaissances morphologiques, connaissances phonémiques et performances en lecture. Chaque variable est représentée par plusieurs indicateurs ou variables observées et par une erreur de mesure. Ce modèle a été construit dans le but de tester les effets directs et indirects des connaissances phonologiques et morphologiques à T1, considérées comme des variables indépendantes, sur la performance en lecture à T3 considérée quant à elle comme la variable dépendante.

Le modèle testé postule que les connaissances phonologiques et morphologiques à T1 ont un pouvoir prédictif important sur les performances en lecture à T3. Les connaissances phonémiques à T2 jouent un rôle médiateur dans cette relation. Dans un premier temps, une analyse factorielle confirmatoire a été réalisée afin de confirmer la structure du modèle. L'objectif est de vérifier que le modèle théorique proposé offre une bonne représentation des phénomènes qu'il est censé représenter. Le test de la validité du modèle de mesure reliant les indicateurs (variables observées) aux construits (variables latentes) a révélé un bon ajustement du modèle théorique aux données avec des indices d'ajustement et une fiabilité satisfaisants. Le modèle, ayant été correctement ajusté aux données de cette recherche, nous avons pu passer à l'étude des liaisons entre ces différents construits ou variables latentes. Les résultats des coefficients des pistes montrent que les connaissances phonologiques et morphologiques ont un effet significatif sur les performances en lecture. Cet effet est direct pour les connaissances morphologiques (β=.27) et indirect pour les connaissances phonologiques

(β=.176) puisqu'il est partiellement médiatisé par les connaissances phonémiques. Ces deux dimensions permettent d'expliquer 46% de la variance. Nous pouvons conclure qu'au cours des années préscolaires, les compétences phonologiques et morphologiques contribuent, bien qu'avec des intensités différentes, significativement et de façon positive à l'explication de la variance dans les performances en lecture en 2ème année. Les connaissances phonémiques à T2 jouent un rôle médiateur dans les effets des connaissances phonologiques sur le niveau de lecture.

Les résultats font aussi ressortir que les connaissances phonologiques à T1 influencent positivement les connaissances phonémiques à T2 (β=.36) et permettent d'expliquer 15% de la variance. Les connaissances phonémiques à T2 semblent affecter, à leur tour, positivement les performances en lecture à T3 (β=.492). En ce qui concerne les relations entre variables indépendantes, les résultats indiquent une corrélation forte et significative entre les connaissances phonologiques et morphologiques, mettant en exergue un chevauchement de ces deux construits.

Le test et l'estimation du modèle nous a permis d'accepter la structure du modèle global et de confirmer notre deuxième hypothèse. En effet, l'ensemble des résultats obtenus, allant des indices d'ajustement du modèle, au test de significativité des paramètres du modèle (coefficients de régression standardisés et C.R associés) confirme que le modèle testé est acceptable. Il est intéressant de constater que ce modèle rend compte de 46% de la variance de la performance en lecture.

Les résultats obtenus à partir du MES confirment, au moins en partie, les résultats des recherches effectuées sur différentes langues (anglais, chinois) et qui ont adopté cette approche (Storch & Witehurst, 2002 ; Muter & al., 2004 ; Leong & al., 2005). Storch & Whitehurst (2002) obtiennent des résultats consistants avec la présente recherche en mettant en évidence des relations similaires entre les connaissances phonologiques précoces et la performance en lecture au début de la scolarité. Les chercheurs trouvent que les habiletés langagières orales n'apportent pas une contribution directe aux capacités de lecture en 1ère et 2ème année. Toutefois, la relation entre les mesures préscolaires du langage oral et la lecture est médiatisée par les habiletés de recodage phonologique.

Le test du modèle a donné des résultats acceptables. Toutefois, il est nécessaire de rappeler qu'un tout autre modèle pourrait également correspondre aux données (on ne rejette pas H0).

La question de la respécification du modèle se pose lorsque l'on cherche à améliorer son ajustement aux données empiriques. Le processus de respécification doit être conduit avec prudence en s'appuyant sur une justification théorique. Par ailleur, au lieu d'adopter une modélisation structurelle et travailler avec des variables latentes, on peut aussi construire un modèle élaboré uniquement à partir de variables observées, on parle dans ce cas de modèle en pistes causales.

Dans le chapitre suivant nous aborderons le rôle et l'apport des évaluations des comportements et compétences scolaires réalisées par les enseignants de CP dans la prédiction du niveau ultérieur en lecture.

Chapitre 4 : Evaluation des enseignants et prédiction de la réussite ultérieure en lecture

Dans le second chapitre de cette partie, nous avons essayé d'étudier la validité prédictive globale des compétences cognitives et langagières précoces chez des enfants suivis de la classe préparatoire à la 2ème année primaire. Toutefois, pour mieux prédire les performances des élèves, il peut s'avérer utile de ne pas limiter les évaluations aux seuls domaines cognitif et langagier, mais de disposer également d'une évaluation des comportements et des compétences des élèves en classe au moyen d'un questionnaire renseigné par les enseignants. Ce type d'évaluation semble ainsi plus proche de la réalité du fonctionnement de l'élève en classe que les évaluations langagières et cognitives standardisées. L'objectif de ce chapitre est de déterminer la capacité d'un questionnaire d'évaluation des comportements et des compétences scolaires renseignés par les enseignants en classe préparatoire (CP) à prédire les performances en lecture de ces mêmes élèves en fin de 1ère et de 2ème années du primaire.

1. Méthode

1.1. Sujets

Treize enseignants de classe préparatoire ont participé à cette étude. Les enseignants étaient tous de sexe féminin et sont âgés de 35 ans en moyenne (écart-type : 9,3). Six d'entre elles avaient une expérience d'au moins 10 ans dans l'enseignement de base, quatre avaient une expérience de 4 ans en moyenne et trois étaient récemment recrutés pour enseigner en classe préparatoire.

1.2. Matériel

L'évaluation des comportements et des compétences scolaires a été réalisée à partir d'un questionnaire[86] administré aux enseignants de classe préparatoire (Florin, Guimard & Nocus, 2008 ; 2002). Les domaines explorés relèvent du langage, du vocabulaire, de la motricité, de la mémoire, etc. Nous avons demandé également un pronostic pour les deux années suivantes en ce qui concerne la réussite en lecture.

[86] Ce questionnaire a été conçu par Florin, Guimard & Nocus (2002 ; 2008) au sein du laboratoire «Education, Cognition, Développement » (Labécd) de l'université de Nantes. En accès libre sur le site : www.lettres.univ-nantes.fr/labecd/index.htm

Le questionnaire présenté aux enseignants de CP est composé de vingt questions appréhendant trois grandes dimensions : les compétences langagières, les comportements des enfants en classe et leurs compétences scolaires. Les compétences langagières sont examinées par les items suivants : niveau de participation verbale en classe (Q1), quantité de participation verbale en classe (Q2), niveau en vocabulaire (Q3), en syntaxe (Q4), en prononciation (Q5) et en compréhension (Q6). Huit items rendent compte de certains aspects comportementaux des enfants en classe : niveau d'attention (Q7), d'autonomie (Q8), rapidité dans l'exécution d'une tâche (Q9), présence de signes de fatigue (Q11), suivi du rythme de la classe (Q12), capacité d'organisation dans les tâches scolaires (Q13), niveau de confiance en soi (Q14) et respect des règles de vie en classe (Q15). Cinq items appréhendent les compétences scolaires ou transversales : niveau en mémorisation (Q10), dans la maîtrise des gestes (Q16), dans les activités d'éducation physique et sportive (Q17), en raisonnement (Q18) et dans l'approche du nombre (Q19). La dernière question est un pronostic de l'enseignant sur l'apprentissage de la lecture l'année suivante (Q20)[87]. Les réponses se distribuent sur une échelle en 3 points correspondant à des niveaux différents (fort, moyen, faible). Les enseignants avaient pour tâche de choisir pour chaque question, la modalité de réponse qui leur semble mieux décrire les compétences scolaires de l'enfant. Un score global est calculé correspondant à la somme des notes obtenues aux 20 items (minimum : 20 ; maximum : 60).

1.3. Procédure

En fin de la classe préparatoire (Mai-juin 2005), nous avons administrés le questionnaire d'évaluation des comportements et compétences scolaires que nous avons traduit en langue arabe aux enseignants (un exemplaire pour chaque élève). Nous avons expliqué aux enseignants les objectifs de la rechreche et qu'il s'agit de noter le comportement le plus habituel de l'enfant au cours de l'année scolaire. L'enseignant indique la modalité de réponse qui lui semble décrire le mieux les comportements de l'enfant. Nous avons insisté sur le fait qu'il est important de répondre à toutes les questions posées et de choisir une des trois modalités de réponse. Le temps nécessaire pour remplir le questionnaire est en moyenne de 3 minutes.

[87] Dans la version française, le pronostic concerne la réussite de l'enfant dans les apprentissages scolaires en général. Dans le cadre de cette recherche, cet item a été légèrement modifié (le pronostic concerne les performances ultérieures en lecture).

2. Résultats

2.1. Caractéristiques du questionnaire renseigné par les enseignants

2.1.1. Corrélations inter-items

La matrice des corrélations présentée dans le tableau 44 indique que les items du questionnaire aux enseignants sont tous fortement corrélés entre eux. Tous les coefficients de corrélations sont significatifs à p<.01. En effet, sur les 190 coefficients calculés, seuls 20 sont inférieurs à .50 dont 14 concernent l'item 11 (fatigue).

Tableau 44 : Matrice des corrélations entre les items du questionnaire

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q11**	**Q12**	**Q13**	**Q14**	**Q15**	**Q16**	**Q17**	**Q18**	**Q19**	**Q20**
1	1																			
2	.75	1																		
3	.73	.74	1																	
4	.67	.63	.82	1																
5	.57	.61	.78	.75	1															
6	.65	.68	.72	.63	.61	1														
7	.62	.62	.67	.61	.63	.64	1													
8	.70	.64	.71	.63	.59	.63	.70	1												
9	.65	.54	.60	.57	.47	.50	.54	.61	1											
10	.67	.68	.72	.65	.68	.65	.63	.65	.59	1										
11	.42	.42	.48	.42	.35	.47	.39	.45	.57	.48	1									
12	.72	.73	.73	.62	.60	.67	.77	.71	.71	.70	.51	1								
13	.66	.63	.55	.63	.52	.61	.60	.61	.68	.65	.53	.79	1							
14	.71	.64	.68	.61	.56	.65	.63	.59	.62	.66	.48	.71	.65	1						
15	.42	.53	.51	.49	.54	.51	.69	.56	.41	.52	.32	.63	.50	.48	1					
16	.72	.62	.69	.69	.58	,65	.59	.67	.59	.74	.48	.68	.70	.67	.51	1				
17	.61	.65	.76	.69	.71	.69	.59	.62	.54	.73	.50	.70	.68	.64	.53	.76	1			
18	.67	.64	.67	.64	.52	.70	.56	.63	.66	.64	.49	.76	.79	.65	.50	.76	.73	1		
19	.67	.68	.77	.67	.63	.66	.62	.65	.63	.68	.52	.76	.71	.69	.58	.73	.74	.78	1	
20	.70	.68	.77	.75	.70	.68	.67	.68	.66	.71	.44	.79	.70	.68	.56	.73	.71	.72	.77	1

Par ailleurs, les plus faibles coefficients de corrélations sont obtenus pour ce même item, le minimum est relevé à .32. L'analyse des corrélations inter-items montre également que la valeur de la corrélation de l'item 15 (respect des règles) avec le reste des items correspond à

un niveau assez faible. Cet item enregistre plus de cinq corrélations inférieures au seuil de .50 avec les autres items du questionnaire.

2.1.2. Analyse de la fiabilité

Afin de vérifier la qualité de cohérence interne du questionnaire, nous avons eu recours au coefficient « Alpha de Cronbach » (α de Cronbach). Les principales caractéristiques du questionnaire enseignant sont résumées dans les tableaux ci-dessous.

Tableau 45 : Statistique descriptive du questionnaire

Nombre d'items	**Score moyen**	**Ecart-type**	**Corrélation moyenne inter-item**	**Minimum/ Maximum**	**Ecart**	**Alpha de Cronbach**
20	2.49	.68	.63	.32 / .82	.50	.97

Globalement, la qualité de cohérence interne du questionnaire telles que mesurées par le coefficient α de Cronbach semble bonne (Alpha = .97). Cependant, ce coefficient est une mesure sensible au nombre d'items dans le questionnaire et sa valeur peut être artificiellement accrue par l'ajout d'items (Roussel & Wacheux, 2006). Ainsi, si l'on fait l'hypothèse que chaque item a un impact positif sur la fiabilité, le fait de retirer un item doit diminuer sa fiabilité. Or, comme le montre le tableau 46, tous les items du questionnaire contribuent à sa fiabilité.

Tableau 46: Valeurs du coefficient Alpha en retirant un item

Item	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
α sans l'item	.97	.97	.96	.97	.97	.97	.97	.97	.97	.97
Item	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
α sans l'item	.97	.96	.97	.97	.97	.96	.97	.96	.96	.96

2.1.3. Analyse factorielle confirmatoire (AFC)

La démarche méthodologique adoptée s'inspire de Roussel & al. (2002) qui soulignent l'intérêt des méthodes d'équations structurelles pour tester la fiabilité et la validité des échelles de mesure.

Nous avons eu recours à l'analyse factorielle confirmatoire afin de tester la validité covergente du questionnaire. La démarche suivie s'articule autour des quatre étapes de spécification du modèle, d'identification, d'estimation et de test de l'ajustement. Le modèle de mesure spécifié dans notre démarche est simple et directement issu de l'approche théorique. En effet, les trois dimensions de la capacité du questionnaire ont été clairement identifiées lors de recherches précédentes (Florin, Guimard & Nocus, 2002 ; Guimard & Florin, 2007 ; 2001). La méthode d'estimation utilisée est celle du maximum de vraisemblance.

Le modèle de mesure qui va être testé comprend trois variables latentes : compétences langagières, comportements en classe et compétences scolaires. Chaque variable latente est appréciée par des variables observées, lesquelles sont les scores moyens aux différents items du questionnaire. L'AFC permet de tester la validité de de chaque dimension du questionnaire.

2.1.3.1. Analyse de la validité

La validité convergente se vérifie à partir des contributions factorielles qui doivent être statistiquement significatives pour chaque dimension. Elle peut être vérifiée si et seulement si le test t (C.R.) associé à chacune des contributions factorielles est supérieur à 1,96 au niveau de signification de 5%. Cette condition est vérifiée pour le construit « compétence langagières » (tableau 47).

Tableau 47 : Test de validité convergente du construit « compétences langagières »

Poids factoriels standardisés	Estimation	C.R.
Q6 <--- Compétences langagières	.801	11.825*
Q5 <--- Compétences langagières	.803	12.072*
Q4 <--- Compétences langagières	.852	13.133*
Q3 <--- Compétences langagières	.925	14.841*
Q2 <--- Compétences langagières	.819	12.402*
Q1 <--- Compétences langagières	.819	12.415*

*p<.05

Ce sont les items 3 (vocabulaire) et 4 (syntaxe) qui apportent les meilleures cobtributions factorielles (respectivement λ =.925 et .852 ; p<.05).

Le construit « comportements en classe » présente également une bonne validité convergente (tableau 48). En effet, le test t associé à chacune des contributions factorielles est significatif

(> 1,96) pour tous les items, avec une supériorité des items 12 (suivi du rythme) et 13 (organisation) (respectivement λ =.917 et .840 ; p<.05).

Tableau 48 : Test de validité convergente du construit « comportement en classe »

Poids factoriels standardisés	Estimation	C.R.
Q15<--- Comportements en classe	.666	9.204*
Q14 <--- Comportements en classe	.796	9.419*
Q13 <--- Comportements en classe	.840	9.856*
Q12 <--- Comportements en classe	.917	10.594*
Q11 <--- Comportements en classe	.589	7.212*
Q9 <--- Comportements en classe	.764	9.090*
Q8 <--- Comportements en classe	.797	9.433*
Q7 <--- Comportements en classe	.802	9.478*

*p<.05

Cette condition est vérifiée également pour le construit « compétences scolaires » puisque les valeurs de t sont toutes significatives au seuil de 5%, condition nécessaire à la validité convergente (voir tableau 49). Tous les items apportent des contributions factorielles non négligeables à la variance (λ supérieur à .80 ; p<.05).

Tableau 49: Test de validité convergente du construit « compétences scolaire »

Poids factoriels standardisés	Estimation	C.R.
Q19 <--- Compétences scolaires	.875	15.825*
Q18 <--- Compétences scolaires	.859	15.562*
Q17 <--- Compétences scolaires	.863	15.728*
Q16 <--- Compétences scolaires	.862	15.697*
Q10 <--- Compétences scolaires	.828	14.522*

*p<.05

L'ensemble des Ratios Critiques (C.R.) pour les trois dimensions est significatif, ce qui suggère que toutes nos variables observables contribuent à l'explication de la variable latente.

Un dernier critère examiné est celui des corrélations entre les variables latentes. Les résultats présentés dans le tableau 50 montrent que toutes les corrélations entre les trois variables latentes sont positives et significatives (r>.90).

Tableau 50 : Corrélations entre les variables latentes

	Estimation
Compétences langagières - Comportements en classe	.91
Compétences langagières - Compétences scolaires	.93
Comportements en classe - Compétences scolaires	.94

Les fortes corrélations relevées entre les variables latentes soulignent le chevauchement de ces trois construits. Les différents résultats de l'analyse factorielle confirmatoire sont présentés dans la figure 12.

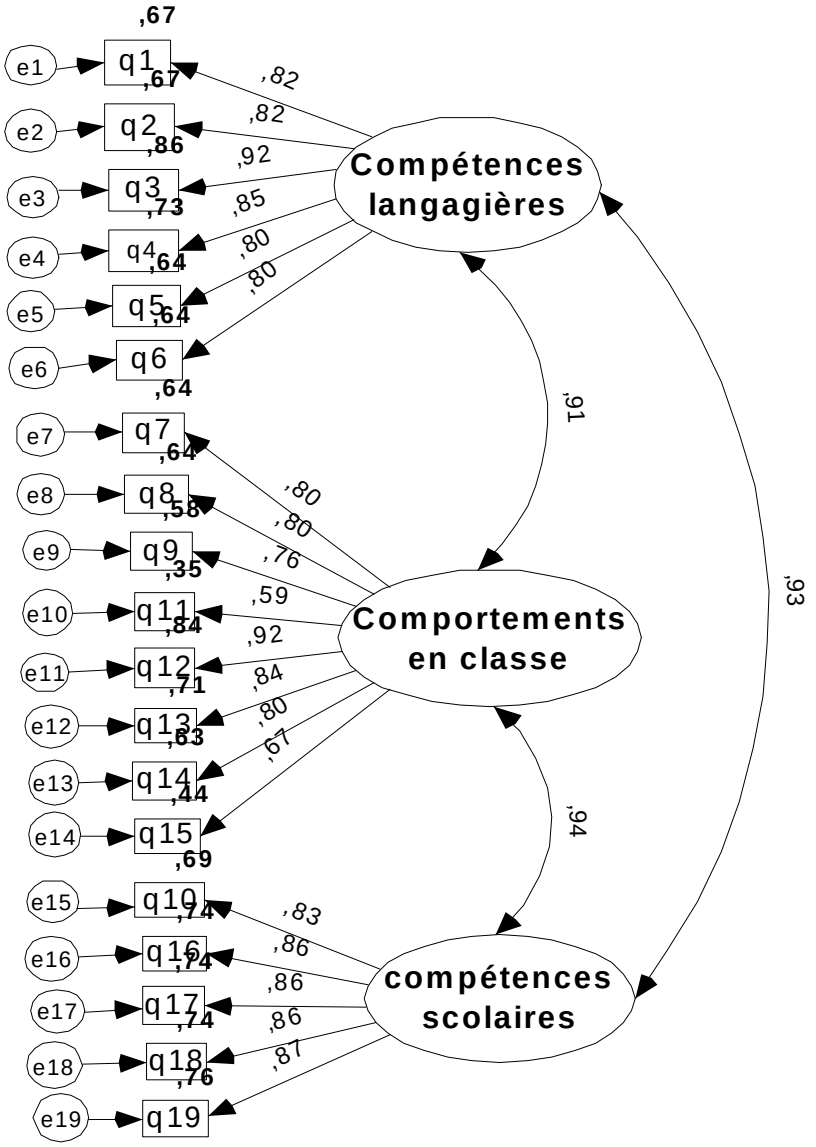


Figure 12 : Résultats de l'analyse factorielle confirmatoire

2.1.3.2. Evaluation de la qualité d'ajustement

L'évaluation des critères du bon ajustement est la phase essentielle de l'AFC avec l'analyse des contributions factorielles. Parmi les indices pertinents dans le cadre d'une AFC, Roussel & al. (2002) retiennent quatre indices absolus (χ2, RMSEA, GFI et AGFI), 3 indices incrémentaux (NFI, IFI, CFI) et un indice de parcimonie (χ2/dl). Le logiciel AMOS nous fournit les résultats suivants :

Tableau 51 : Indices d'ajustement du modèle

Indice	χ^2	dl	χ2/dl	GFI	AGFI	NFI	CFI	IFI	RMSEA
Valeur	415.4	149	2.788	.912	.886	.942	.914	.952	.005
Seuil	ns	-	<5	>.90	>.90	>.90	>.90	>.90	<.08

D'après le tableau 51, nous constatons que le modèle présente une bonne qualité d'ajustement. Le rapport χ2/dl est inférieur à 5. Les indicateurs de qualité globale de l'ajustement (GFI, AGFI, NFI, CFI, TLI) sont supérieurs à .90. Le RMSEA a une valeur acceptable (.005). Le modèle testé présente des indices d'ajustement très satisfaisants et proches des normes empiriques établies.

L'examen des coefficients de corrélation multiples (R^2) montre qu'ils sont acceptables dans la grande majorité des cas. L'analyse des coefficients de régression de chaque indicateur sur la variable latente montre que tous les items apportent un pouvoir explicatif non négligeable à la variance (R^2 proche de .70), à l'exception des items Q11 et Q15. Le pouvoir explicatif de ces deux items est relativement faible par rapport aux autres items (R^2=.34 pour l'item 11 et R^2=.44 pour l'item 15). Ces résultats attestent d'un important pouvoir explicatif et prouvent que les variables exogènes sont bien expliquées par les variables endogènes (items) qui leurs sont associées. L'ensemble des indices examinés permet de mettre en évidence une bonne qualité d'ajustement du modèle théorique.

2.2. Relation prédictive entre les évaluations des enseignants en CP (T1) et les performances en lecture en 1ère année (T2)

L'objectif est ici d'évaluer la valeur prédictive de l'évaluation réalisée par les enseignants sur le niveau de lecture en 1ère année évalué par l'épreuve d'identification de mots écrits (IME1).

2.2.1. Analyse des corrélations

Les corrélations entre les différents items du questionnaire renseigné par les enseignants de CP et le niveau de lecture en 1ère année sont présentées dans le tableau 52. Nous constatons que tous les items du questionnaire sont corrélés significativement avec les performances en identification de mots. Il apparait que 6 corrélations parmi les 20 atteignent un niveau supérieur ou égal à .40.

Tableau 52 : Corrélations entre les items du questionnaire à T1 et les performances en identification de mots à T2 (IME1)

Items du questionnaire	**IME1**
Q1 Quantité de participation en classe	**.40****
Q2 Qualité de participation en classe	.36*
Q3 Vocabulaire	.36*
Q4 Syntaxe	.34*
Q5 Prononciation	.35*
Q6 Compréhension	.29*
Q7 Attention	.38*
Q8 Autonomie	.32*
Q9 Rapidité	.30*
Q10 Mémoire	.38*
Q11 Fatigue	.26*
Q12 Suivi du rythme	**.44****
Q13 Organisation	**.46****
Q14 Confiance en soi	.38*
Q15 Respect des règles	**.42****
Q16 Maîtrise de gestes	.36*
Q17 Activités EPS	.38*
Q18 Raisonnement	.37*
Q19 Approche du nombre	**.43****
Q20 Pronostic	**.42****
Score global	**.46****

* s à p<.05 ** s à p<.01

En gras : les corrélations les plus fortes.

Les corrélations les plus fortes sont trouvées pour l'organisation dans les tâches scolaires (r=.46 ; p<.01), suivi du rythme de la classe (r=.44 ; p<.01), l'approche du nombre (r=.43 ;

p<.01), respect des règles de vie en classe et pronostic de réussite ultérieure en lecture (r=.42 ; p<.01) et la quantité de participation en classe (r=.40 ; p<.01). Le score global au questionnaire représente la dimension la plus fortement corrélée à l'épreuve d'identification de mots (r=.46 ; p<.01).

Ainsi, les évaluations réalisées par les enseignants en CP sont dans l'ensemble fortement liées aux performances en lecture. Les corrélations les plus élevées étant obtenues avec le score global.

2.2.2. Analyse de régressions multiples

Une analyse de régression pas à pas a été conduite afin de déterminer dans quelle mesure les évaluations des comportements et compétences scolaires par les enseignants prédisent, de façon indépendante, les performances en lecture à T2. Les items du questionnaire sont considérés comme des variables explicatives. Les scores obtenus à l'épreuve d'identification de mots (IME1) est la variable dépendante.

Tableau 53 : Analyse de régression pas à pas entre les items du questionnaire à T1 et les performances en identification de mots à T2 (IME1)

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Pas 1 : Q13 Organisation	.255	.202	.197	42.678
Pas 2 : Q15 Respect des règles	.245	.235	.226	25.852
Pas 3 : Q1 Quantité de participation	.298	.257	.244	19.276
Pas 4 : Q8 Autonomie	.205	.274	.257	15.700

Les résultats présentés dans le tableau 53 indiquent que les variables les plus explicatives du niveau ultérieur en identification de mots sont : la capacité d'organisation dans les tâches scolaires (Q13), le respect des règles de vie en classe (Q15), la quantité de participation en classe (Q1) et le niveau d'autonomie (Q8). Ces quatre variables expliquent à elles seules 25% de la variance totale. Les deux variables comportementales (capacité d'organisation et respect des règles) expliquent la part la plus grande de variance des performances en identification de mots (22%). La quantité de participation et l'autonomie ajoutent respectivement 2% et 1% de variance supplémentaire.

L'ensemble de ces résultats a permis de mettre en évidence des liens assez forts entre les évaluations des compétences et comportements scolaires réalisées par les enseignants de CP et les performances en lecture en 1ère année.

Afin de comparer le pouvoir prédictif du questionnaire aux enseignants et celui des compétences langagières mesurées en CP, une analyse de régression a été réalisée à partir d'un modèle incluant comme variables indépendantes le score global au questionnaire et les mesures langagières les plus prédictives de CP (connaissances des lettres, connaissances phonologiques et morphologiques). Les scores à l'épreuve IME1 constituent la variable dépendante. Les résultats indiquent que l'évaluation réalisée par les enseignants n'entre pas dans les équations de régression et ne contribuent pas à l'explication de la variance totale. La part explicative du questionnaire reste modeste comparée à celle apportée par les compétences langagières.

2.3. Relation prédictive entre les évaluations des enseignants en CP (T1) et les compétences en lecture en 2ème année (T3)

2.3.1. Analyse des corrélations

Le tableau 54 présente les corrélations entre les items du questionnaire renseigné par les enseignants en CP et les performances en lecture en 2ème année. Les corrélations les plus fortes avec les performances à l'épreuve IME2 sont observées pour l'organisation dans les tâches scolaires ($r=.44$; $p<.01$), le niveau de vocabulaire et le suivi du rythme ($r=.43$; $p<.01$), la syntaxe ($r=.43$; $p<.01$), la confiance en soi, le pronostic ($r=.41$; $p<.01$) et la mémoire ($r=.40$; $p<.01$).

Les corrélations entre les items du questionnaire et le Timé2 sont plus nombreuses et plus élevées puisque 15 corrélations parmi les 20 atteignent un niveau élevé, supérieur ou égal à .40. Les liaisons les plus significatives concernent l'organisation ($r=.52$; $p<.01$), la confiance en soi et la quantité de participation en classe ($r=.51$; $p<.01$), le suivi du rythme ($r=.48$; $p<.01$), le vocabulaire ($r=.47$; $p<.01$), la mémoire ($r=.46$; $p<.01$), le pronostic et la qualité de participation en classe ($r=.45$; $p<.01$).

S'agissant des liens entre les évaluations des enseignants et les performances en lecture-compréhension, nous constatons des corrélations moins nombreuses et moins fortes que précédemment. Les corrélations les plus significatives sont trouvées pour la quantité de participation en classe ($r=.39$; $p<.01$) et le pronostic ($r=.36$; $p<.01$). Le score global au

questionnaire apparait comme l'indicateur le plus fortement corrélé au niveau ultérieur en lecture, quelle que soit l'épreuve de lecture utilisée (r=.47 ; p<.01 avec IME2 ; r=.53 ; p<.01 avec Timé2 ; r=.39 ; p<.05 avec la compréhension écrite).

Tableau 54 : Corrélations entre les items du questionnaire aux enseignants et les performances en lecture à T3

Items du questionnaire	**IME2**	**Timé2**	**Cph écrite**
Q1 Quantité de participation en classe	**.42****	**.51****	**.39***
Q2 Qualité de participation en classe	.38*	**.45****	.32*
Q3 Vocabulaire	**.43****	**.47****	.34*
Q4 Syntaxe	**.42****	**.43****	.29*
Q5 Prononciation	.38*	.37*	.31*
Q6 Compréhension	.37*	**.44****	.30*
Q7 Attention	.36*	.38*	.31*
Q8 Autonomie	.29*	**.40****	.31*
Q9 Rapidité	.36*	**.42****	.30*
Q10 Mémoire	**.40***	**.46****	.34*
Q11 Fatigue	.25*	.23*	.19
Q12 Suivi du rythme	**.43****	**.48****	.32*
Q13 Organisation	**.44****	**.52****	.35*
Q14 Confiance en soi	**.41***	**.51****	**.36***
Q15 Respect des règles	.37*	.36*	.25*
Q16 Maîtrise des gestes	.32*	**.44****	.33*
Q17 Activités EPS	.34*	.38*	.24*
Q18 Raisonnement	.37*	**.48****	.31*
Q19 Approche du nombre	.38*	**.43****	.29*
Q20 Pronostic	**.41****	**.45****	**.36***
Score global	**.47****	**.53****	**.39***

* s à p<.05 ** s à P<.01

En gras : les corrélations les plus fortes.

2.3.2. Analyses de régressions multiples

Des analyses de régression pas à pas ont été conduites afin d'examiner le pouvoir prédictif des items du questionnaire renseigné par les enseignants (variable indépendantes) sur les performances en lecture à T3 évaluées par l'épreuve d'identification de mot (IME2), le Timé2 et la compréhension écrite (variables dépendantes).

2.3.2.1. Analyse de régressions sur les performances d'identification de mots à T3

Les résultats obtenus avec l'épreuve IME2 (tableau 55) montrent que le niveau de vocabulaire estimé par les enseignants explique la plus forte part de variance expliquée (18%). La seconde variable explicative est comportementale : il s'agit de la capacité d'organisation dans les tâches scolaires qui ajoute 4% de variance supplémentaire.

Tableau 55 : Analyses de régression pas à pas entre les items du questionnaire et les performances à l'épreuve IME2 à T3

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Pas 1 : Q3 Vocabulaire	.266	.191	.186	39.907
Pas 2 : Q13 Organisation	.262	.231	.222	25.187

L'analyse de régression réalisée entre les items du questionnaire et les performances au Timé2 (tableau 56) indique que les variables les plus explicatives du niveau de lecture en 2ème année sont comportementales : l'organisation et la confiance en soi qui expliquent à elles seules 31% de la variance en lecture. La quantité de participation en classe ajoute une part explicative modeste (1% de variance supplémentaire).

Tableau 56 : Analyses de régression pas à pas entre les items du questionnaire et les performances au Timé2 à T3

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Pas 1 : Q13 Organisation	.259	.271	.267	62.827
Pas 2 : Q14 Confiance en soi	.200	.321	.313	39.642
Pas 3 : Q1 Quantité de participation	.196	.337	.325	28.298

Deux analyses de régression analogues ont été réalisées afin de comparer le pouvoir prédictif du questionnaire aux enseignants et celui des mesures langagières réalisées en CP sur le niveau de lecture en 2ème année (voir tableau 57). Lorsqu'on introduit le score global au questionnaire et les mesures langagières de CP (variables indépendantes), on constate que la performance en identification de mots est expliquée principalement par le score global au questionnaire (tableau 50). Ainsi, contrairement aux résultats obtenus en 1ère année, les évaluations réalisées par les enseignants en CP expliquent la plus forte part de variance explicative, quelle que soit l'épreuve de lecture utilisée (22% à l'épreuve IME2 et 28% au

Timé2) et constituent le meilleur prédicteur des performances en lecture des élèves en 2ème année.

Tableau 57 : Analyse de régression pas à pas entre les scores globaux au questionnaire et aux mesures langagières à T1 et les scores en lecture à T3

Variables expliquées			
IME2		**Timé2 arabe**	
Variables prédictives	*R2*	*Variables prédictives*	*R2*
SG Questionnaire	22%	SG Questionnaire	28%
Connaissance des lettres	27%	Connaissance des lettres	38%
Segmentation syllabique	30%	Détection intrus syllabique	42%
Comptage syllabique	32%		
Variance totale expliquée	32%	Variance totale expliquée	42%

Dans l'ensemble, ces résultats montrent que le score global au questionnaire constitue systématiquement le meilleur prédicteur des performances en identification de mots en 2ème année quelle que soit la variable dépendante à expliquer. Cependant, il apporte une part explicative plus importante au Timé2.

2.3.2.2. Analyse de régressions sur les performances en lecture-compréhension à T3

L'analyse de régression réalisée entre les items du questionnaire et l'épreuve de compréhension écrite montre que la quantité de participation en classe est la seule variable prédictive : elle explique 16% de la variance totale (voir tableau 58).

Tableau 58 : Analyse de régression pas à pas entre les items du questionnaire et les performances en lecture-compréhension à T3

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
Q1 Quantité de participation en classe	.399	.159	.154	32.041

2.4. Les liens prédictifs entre les évaluations des enseignants et le niveau ultérieur en lecture : proposition d'un modèle structurel

Afin d'étudier les les relations prédictives entre les évaluations des enseignants et les performances en lecture et de vérifier les hypothèses émises dans le cadre de ce travail, la méthodologie des modèles en équations structurelles a été adoptée. Chacune des trois

dimensions du questionnaire est spécifiée comme variable latente (figurée par une forme ovale) : compétences langagières, comportements en classe et compétences scolaires. Elle représente un facteur commun de plusieurs variables (représentées par des rectangles ou carrés) correspondant aux items du questionnaire. Ainsi, nous allons essayer de tester la relation entre les différentes dimensions du questionnaire et les performances en lecture à T3.

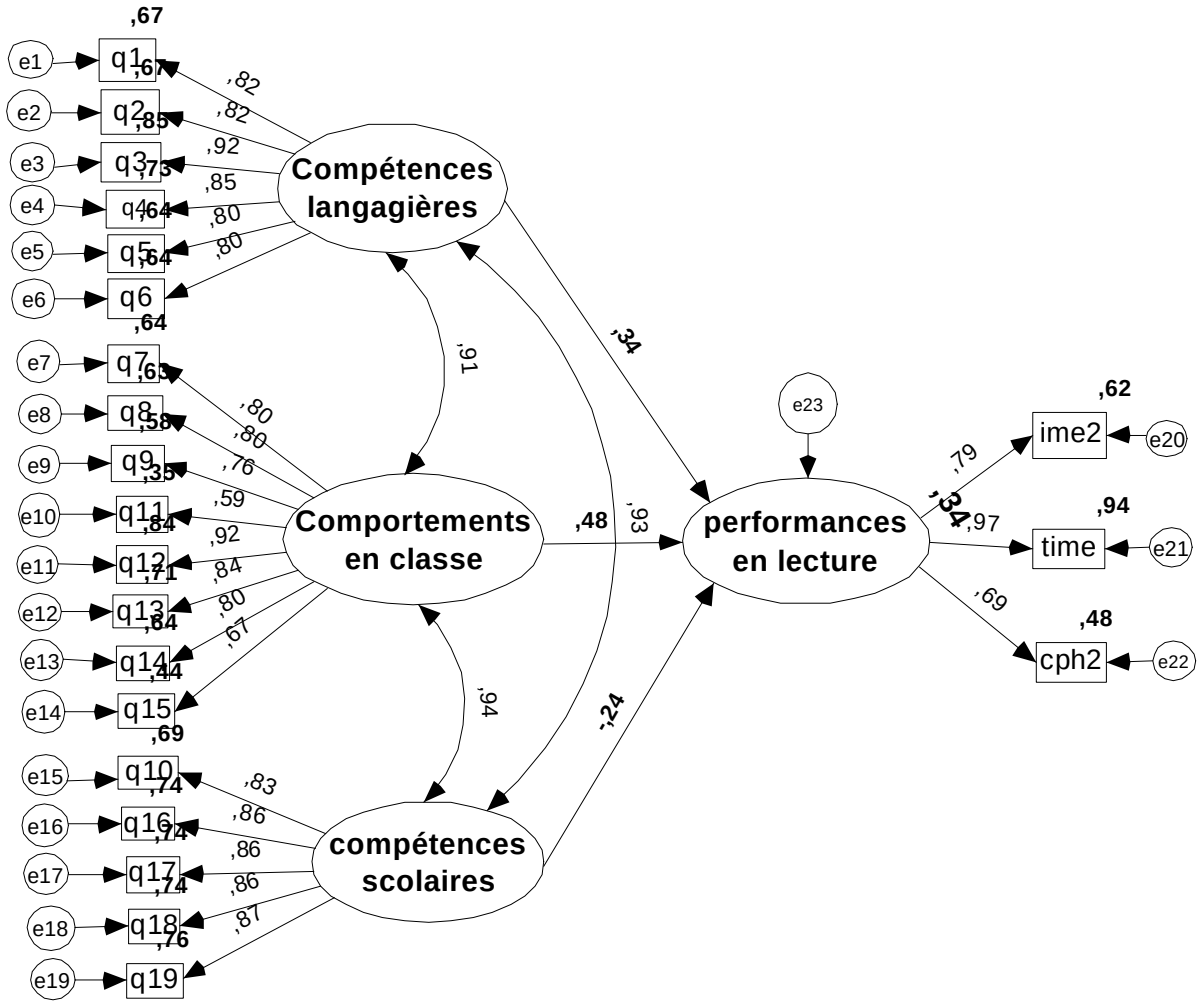


Figure 13 : Modèle d'équations structurelles des relationsprédictives entre les dimensions du questionnaire et les performances en lecture à T3[88]

Les trois dimensions peuvent être ainsi spécifiées comme variables latentes explicatives dont les effets sur la performance en lecture (variable latente à expliquer) peuvent être évalués par les coefficients de causalité qui relient les construits de notre modèle à tester. La question ici posée est donc de savoir comment le modèle théorique peut rendre compte de l'existence

[88] Q1 : quantité de participation en classe ; Q2 : qualité de participation en classe; Q3 : vocabulaire ; Q4 : syntaxe ; Q5 : prononciation ; Q6 : compréhension ; Q7 : attention ; Q8 : autonomie ; Q9 : rapidité ; Q10 : mémoire ; Q11 : fatigue ; Q12 : suivi du rythme ; Q13 : organisation ; Q14 : confiance en soi ; Q15 : respect des règles ; Q16 : maîtrise des gestes ; Q17 : activités EPS ; Q18 : raisonnement ; Q19 : approche du nombre.

d'une relation prédictive entre ces trois dimensions du questionnaire et les performances en lecture (figure 13).

2.4.1. Evaluation de la qualité d'ajustement

Après avoir testé le modèle à l'aide des équations structurelles, il ressort que les conditions d'ajustement du modèle testé aux données sont globalement respectées. Les résultats sont présentés dans le tableau 59. Le $\chi 2/dl$ est inférieur à 5 avec une valeur de 2.32. Le GFI est acceptable avec une valeur de .932 tout comme l'AGFI avec une valeur de .904. Le RMSEA de notre modèle est de .005, preuve d'une excellente adéquation du modèle aux données. Les indices relatifs de comparaison ne sont que moyennement satisfaisants, en particulier le NFI avec une valeur de .863. Le CFI atteint le seuil fixé de .90 avec une valeur de .917, ce qui nous paraît correct. Le NNFI n'est pas supérieur au seuil de .90 mais est très proche avec une valeur de .893. Le IFI prend une valeur de .918 ce qui est satisfaisant. Le modèle testé semble donc assez bien ajusté aux données.

Tableau 59 : Ajustement et pouvoirs explicatifs du modèle

Indices	**$\chi 2$**	**dl**	**$\chi 2/dl$**	**GFI**	**AGFI**	**NFI**	**CFI**	**IFI**	**NNFI**	**RMSEA**
Valeur	471.9	203	2.325	.932	.904	.863	.917	.918	.893	.005
Seuil	ns	-	<5	>.90	>.90	>.90	>.90	>.90	>.90	<.08

Afin de pouvoir confirmer que le modèle reproduit de manière correcte les données, il est nécessaire de compléter cette analyse par la vérification de la significativité des relations entre les variables latentes (Roussel et al., 2002).

2.4.2. Vérification de l'existence de relations de causalité

Après avoir évalué les critères de bon ajustement, l'étape suivante consiste à estimer les paramètres du modèle. Les résultats présentés dans le tableau 60 laissent apparaître que les comportements scolaires à T1 influencent directement la performance en lecture sur la base d'une relation significative (β=.482 ; t=2.041 au seuil de 5%). Les compétences scolaires semblent avoir un effet direct sur la performance en lecture, mais ce coefficient de régression est non significatif (β=-.243 ns). Les résultats montrent également que les compétences langagières n'affectent pas significativement la performance en lecture à T2 (β=.344 ns).

Par ailleurs, l'examen des pourcentages de variance des variables à expliquer prouve que les trois dimensions du questionnaire (variables exogènes) expliquent 34% de la variance en lecture. Ces résultats permettent de dire que la variable endogène est relativement bien expliquée par les variables exogènes qui lui sont associées.

Tableau 60 : Estimation par le maximum de vraisemblance de la significativité des paramètres du modèle

Coefficients de régression standardisés	Estimation	C.R.
Performances en lecture <--- Compétences langagières	.344	1.532 ns
Performances en lecture <--- Comportements en classe	.482	1.961*
Performances en lecture <--- Compétences scolaires	-.243	-.874 ns

* sig à p<.05

3. Discussion

L'objectif principal de ce chapitre était d'étudier la valeur prédictive des évaluations des compétences et comportements scolaires réalisées par les enseignants de CP sur les performances en lecture en 1ère année et en 2ème année du primaire. En effet, nous avons posé l'hypothèse selon laquelle les enseignants de la classe préparatoire auraient un rôle important à jouer dans la prédiction de la réussite dans l'apprentissage de la lecture et dans le repérage des difficultés des élèves. Nous nous sommes interrogés également sur la pertinence de ces évaluations par rapport aux mesures langagières et cognitives utilisés précédemment. Le questionnaire présenté aux enseignants a été conçu comme un outil de dialogue avec les professionnels concernés dans le but de permettre un premier niveau de repérage des difficultés (Florin, Guimard & Nocus, 2008). La valeur prédictive de ce questionnaire a été testée dans plusieurs recherches réalisées de la maternelle à l'école élémentaire (Florin & al., 2002 ; Guimard & Florin, 2007 ; Guimard & al., 2007 ; Guimard & Florin, 2001). Dans la présente recherche, nous avons examiné la valeur prédictive de ce questionnaire, traduit en langue arabe, sur un échantillon d'enfants tunisiens et tenté de confirmer ainsi le poids des évaluations des comportements et des compétences scolaires réalisées par les enseignants sur la réussite en lecture dans le contexte tunisien.

L'analyse des corrélations inter-items indique une liaison forte et significative entre les différents items du questionnaire, ce qui suggère une forte homogénéité du questionnaire. En effet, plus l'homogénéité est élevée, plus l'échelle est fidèle au sens de la cohérence interne

des items, et plus on peut croire que les items mesurent un seul et même construit (Roussel & Wacheux, 2006). La fiabilité du questionnaire aux enseignants a été confirmée grâce à l'utilisation de l'alpha de Cronbach.

Dans le but d'étudier plus finement la structure du questionnaire, une analyse factorielle confirmatoire est réalisée sur l'ensemble des items. Le modèle de mesure étudié était composé de trois variables latentes correspondant aux trois dimensions du questionnaire : comportements en classe, compétences langagières et compétences scolaires. L'AFC a montré que tous les coefficients de saturation entre les items et les variables latentes sont tous positifs et significatifs et a permis de vérifier l'ajustement très acceptable du modèle. La méthode des équations structurelles a permis ainsi de confirmer la fiabilité et la validité du questionnaire enseignant. En revanche, les libellés de certains items nécessitent d'être modifiés, en arabe, pour améliorer les qualités psychométriques du questionnaire, et plus particulièrement les items 11 (fatigue) et 15 (respect des règles). La faible saturation sur la dimension attendue pourrait s'expliquer par une mauvaise traduction de ces items.

Dans le but d'évaluer le poids prédictif des évaluations des enseignants sur les performances ultérieures en lecture, une première série d'analyses a consisté à examiner les corrélations entre les différents items du questionnaire à T1 et les performances à l'épreuve d'identification de mots écrits à T2. Les résultats indiquent que trois variables comportementales sont assez fortement corrélées aux performances en identification de mots (organisation dans les tâches scolaires, suivi du rythme de la classe et respect des règles de vie en classe). Plusieurs items langagiers sont également liés significativement à cette évaluation en lecture, mais de manière un peu plus faible (quantité et qualité de la participation en classe et niveau du vocabulaire). Le score global au questionnaire apparait l'indicateur le plus fortement corrélé à l'épreuve de lecture. L'existence d'un lien significatif entre les évaluations de l'enseignant de CP et la réussite ultérieure en lecture a été confirmée. Ces résultats vont dans le sens de ceux établis antérieurement par Florin (1991), Florin & al. (1999), Guimard & Florin (2001) et Ecalle (2000a) qui soulignent l'intérêt non négligeable des évaluations des enseignants pour effectuer une prédiction globale des performances scolaires (principalement en lecture) à court terme.

L'analyse de régression pas à pas réalisée entre les items du questionnaire et les performances à l'épreuve IME1 montre que trois items comportementaux (l'organisation en classe, le respect des règles et l'autonomie) expliquent une part importante de la variance totale (23%).

Une compétence langagière (quantité de participation en classe) contribue modestement à la variance en lecture (2%). Dans le but de comparer la prédiction réalisée par les enseignants à celle issue des mesures langagières et cognitives proposées en classe préparatoire, nous avons réalisé une analyse de régression en incluant le score global au questionnaire et les mesures langagières de CP, en tant que variables indépendantes, dans le modèle. Les résultats montrent que la part explicative du questionnaire est non significative comparée à celles apportées par les épreuves langagières standardisées. Ces résultats ne sont pas en accord avec ceux obtenus par Guimard & Florin (2007) qui ont trouvé que l'évaluation réalisée par les enseignants en GSM est le meilleur prédicteur global à court terme des performances en lecture des élèves. Le score global au questionnaire en GSM constitue un meilleur prédicteur des performances en lecture au CP que les mesures cognitives et langagières de la BREV (41% de variance expliquée).

En 2ème année primaire, deux autres épreuves de lecture ont été utilisées : un test d'identification de mots écrits inspiré du Timé2 (Ecalle, 2003) et une épreuve de compréhension écrite adaptée de Khomsi (1990). L'objectif est d'analyser l'évolution au cours de la scolarité du caractère prédictif du questionnaire renseigné par les enseignants. Les analyses de corrélation ont permis de mettre en évidence des liens assez forts entre les évaluations réalisées par les enseignants en CP et les épreuves de lecture proposées en 2ème année. Conformément aux études antérieures (Florin & al., 2002 ; Guimard & Florin, 2001 ; 2007), certains items langagiers (quantité et qualité de participation, niveau de vocabulaire et de syntaxe) et comportementaux (suivi du rythme de la classe, organisation et confiance en soi) sont assez fortement corrélés aux scores ultérieurs en lecture. Les corrélations les plus élevées sont obtenues avec le Timé2. Certains items appréhendant les compétences langagières et transversales semblent gagner de leur pouvoir à l'épreuve IME2. C'est particulièrement le cas des items évaluant le niveau en vocabulaire, en syntaxe et en mémoire.

Nous constatons également des relations significatives entre la plupart des items du questionnaire et les performances au Timé2. Quatre d'entre eux atteignent un niveau très élevé : organisation, confiance en soi, quantité de participation en classe et suivi du rythme. Les relations entre les évaluations des enseignants et les performances en compréhension écrite sont moins nombreuses et moins fortes que précédemment. Les meilleures corrélations sont obtenues pour la quantité de participation en classe, la confiance en soi et l'organisation. Les résultats montrent de nouveau que le score global au questionnaire constitue l'indicateur le plus fortement corrélé aux performances ultérieures en lecture quelle que soit l'épreuve de

lecture utilisée. Ces résultats sont assez proches de ceux obtenus par Guimard & Florin (2007 ; 2007) qui montrent que le niveau de confiance en soi et la capacité de s'organiser dans les tâches scolaires estimés par les enseignants en maternelle sont très significativement corrélés aux performances de ces mêmes élèves en lecture et aux évaluations de français ou de mathématique des années plus tard.

Les analyses de régression confirment les résultats des corrélations. Elles indiquent, en effet, que certains aspects comportementaux de CP contribuent pour une part non négligeable à l'explication des performances en lecture (tableau 61). On note en particulier le poids explicatif de l'organisation en classe qui explique 27% de la variance au Timé2 et ajoute 4% de variance supplémentaire à l'épreuve IME2. De même, la confiance en soi apporte un gain de pouvoir explicatif de 5% au Timé2.

Tableau 61 : Récapitulatif des items du questionnaire les plus prédictifs des performances en lecture à T2 et T3

Items du questionnaire			
IME1	**IME2**	**Timé2**	**Cph**
Organisation Respect des règles Quantité participation autonomie	Vocabulaire Organisation	Organisation Confiance en soi Quantité participation	Quantité participation
25% de variance expliquée	**22% de variance expliquée**	**32% de variance expliquée**	**15% de variance expliquée**

Les évaluations des enseignants de CP présentent ainsi une forte stabilité prédictive dans le temps puisque, en 1ère année comme en 2ème année, elles contribuent pour une part non négligeable et constante à l'explication des performances en lecture. Ces résultats vont dans le sens des données de la littérature (Florin & al., 2002 ; Guimard & Florin, 2001 ; 2007 ; Quay & Steele, 1998).

Lorsqu'on introduit les variables langagières de CP, on constate que les performances en lecture sont principalement expliquées par le score global au questionnaire. Il apparaît ainsi que les prédictions globales réalisées à partir du questionnaire sont meilleures que celles obtenues à partir des variables langagières. Le score global au questionnaire constitue le meilleur prédicteur du niveau de lecture en 2ème année et explique la part la plus grande de

variance sur les deux épreuves d'identification de mots (22% à l'épreuve IME, 28% au Timé2). Ces résultats sont en accord avec ceux de Guimard & Florin (2007) qui ont trouvé que les évaluations réalisées par les enseignants en GSM, tout comme les évaluations cognitives et langagières de la BREV, sont assez fortement liées aux performances en lecture en CP. Toutefois, les prédictions globales réalisées à partir du questionnaire sont meilleures que celles obtenues à partir des scores langagier et cognitif à la BREV.

Par ailleur, l'analyse des corrélations entre le score gloabal au questionnaire et les différentes épreuves langagières et cognitives de CP (voir annexe II) montre une liaison forte et significative entre les évaluations des enseignants et les épreuves les plus prédictives des capacités ultérieures en lecture, notamment la connaissance des lettres ($r=.52$; $p<.01$), la suppression syllabique ($r=.52$; $p<.01$), la morpho-syntaxe et le comptage syllabique ($r=.45$; $p<.01$).

Ces résultats permettent donc de soutenir la troisième hypothèse de cette recherche, et révèlent les bonnes qualités psychométriques du questionnaire. La construction d'une échelle de mesure des évaluations des enseignants ouvre des perspectives importantes en termes de recherches sur la prévention des difficultés d'apprentissage scolaire et son efficacité.

Les résultats des analyses en équations structurelles montrent que les évaluations des enseignants relatives aux comportements des élèves en classe sont significativement reliées à la performance des élèves en lecture en 2ème année. Les comportements des élèves en classe évalués par les enseignants constituent ainsi des prédicteurs puissants des performances en lecture. Toutefois, les résultats n'ont pas mis en évidence l'existence d'un lien significatif entre les dimensions relatives aux compétences langagières et aux compétences scolaires et les performances en lecture, ce qui est dû probablement à la liaison étroite entre les variables latentes explicatives. Les fortes corrélations observées entre les trois dimensions du questionnaire ($r>.90$) soulignent le chevauchement de ces construits et l'intérêt de procéder à une AFC de second ordre.

Au terme de ce chapitre, on peut conclure que les évaluations des enseignants constituent, dans le contexte tunisien, une source d'information intéressante dans une perspective de dépistage et de prévention des difficultés scolaires des enfants et plus particulièrement des difficultés dans l'apprentissage de la lecture. Un échange avec l'enseignant fournira des éléments pour que celui-ci mette en oeuvre une prévention précoce par des activités

pédagogiques différenciées et plus adaptées à leurs difficultés. Il peut aider aussi l'enseignant à prendre en compte certaines dimensions des comportements auxquelles il n'aurait pas été spontanément sensible (Guimard & Florin, 2008). Il serait alors possible d'identifier dès la classe préparatoire des enfants « à risque », ce qui nous amène à nous interroger sur la valeur prédictive individuelle du questionnaire et des mesures prédictives de CP.

Chapitre 5 : Valeur prédictive individuelle des compétences langagières et cognitives précoces et des évaluations des enseignants

L'apprentissage de la lecture s'avère être un des enjeux socio-économiques prioritaires pour la période actuelle et à venir. Cependant, certains enfants rencontrent, dès le plus jeune âge, des difficultés dans l'acquisition de la langue écrite. Les difficultés de l'apprentissage de la lecture sont liées à des expériences d'échec scolaire au sein de l'école obligatoire. Le travail de dépistage des sujets « à risque » devrait permettre ainsi de réduire la probabilité d'échec scolaire à travers des actions de nature éducatives ciblées et spécifiques. Il est donc important d'attirer l'attention sur la prévention et l'action précoce qui sont des solutions beaucoup plus efficaces et économiques que des actions plus tardives.

L'objectif de ce chapitre est de mesurer la valeur prédictive individuelle des différentes mesures de CP dans une perspective de prévention. Ainsi, nous essayerons de savoir si les évaluations langagières et cognitives et les évaluations réalisées par les enseignants en CP permettent de prédire individuellement la réussite et les difficultés des élèves en lecture une année et deux années plus tard. En d'autres termes, nous allons essayer de déterminer la proportion des normolecteurs et des enfants en difficulté de lecture correctement détectés par les différentes mesures de CP.

1. Critères de sélection

Afin d'examiner le nombre d'enfants de 1ère année et de 2ème année classés correctement par les prédicteurs de CP et de 1ère année et compte tenu des résultats exposés dans les chapitres précédents, deux scores globaux en langage et dans le domaine cognitif sont calculés. Le score global est obtenu par la somme des scores composites divisée par le nombre d'épreuves. Le score langage à T1 regroupe les résultats aux épreuves de phonologie, de morphosyntaxe et de connaissance des lettres. Le score cognitif à T1 est calculé sur les épreuves d'intelligence, de graphisme, de mémoire, d'attention et de concepts de base. Nous avons retenu également le score global obtenu au questionnaire destiné aux enseignants. Le score langage à T2 regroupe les scores globaux obtenus aux épreuves phonémiques et à l'épreuve de jugement de liens morphologiques. Les enfants ont été répartis dans les catégories

« réussite » et « difficulté » sur les dimensions évaluées en fonction d'un même critère statistique : scores en dessous ou en dessus du 25ème centile.

1.1. Sélection des enfants à partir des prédicteurs de CP et de 1ère année

Le tableau 62 présente les moyennes, les écart-types et les centiles 25 obtenus par les enfants aux différents indicateurs globaux de CP et de 1ère année.

Tableau 62 : Moyennes, écart-types et centiles 25 des performances aux scores globaux langagiers et cognitifs et au questionnaire enseignants

	Moyenne	**Ecart-type**	**Centile 25**
SG langage T1	62	17,8	50
SG cognitif T1	33	7,6	28
SG langage T2	28	6,2	25
SG Questionnaire T1	49	10,6	42

Les élèves considérés en difficulté en langage à T1 ont un score global inférieur ou égal à 50, ce qui correspond à 43 élèves sur 171 (23%). Les autres élèves sont donc considérés en réussite sur ce critère (tableau 63).

Tableau 63 : Enfants prédits en difficulté et en réussite en fonction de leurs scores globaux en langage, dans le domaine cognitif et au questionnaire enseignant

	Prédits en difficulté	**Prédits en réussite**
SG langage T1	43/171 (23%)	128/171 (74%)
SG cognitif T1	46/171 (26%)	125/171 (73%)
SG langage T2	42/171 (24%)	129/171 (75%)
SG Questionnaire T1	43/171 (23%)	128/171 (74%)

Les élèves présentant des difficultés aux épreuves cognitives ont un score inférieur ou égal à 28, soit 46 élèves sur 171 (26%). Ceux présentant des difficultés en langage à T2 ont un score global inférieur ou égal à 25, ce qui correspond à 42 élèves sur 171 (24%). Les élèves

considérés en difficulté par les enseignants obtiennent un score global au questionnaire inférieur ou égal à 42, ce qui correspond à 43 élèves sur 171 (23%).

1.2. Sélection des enfants à partir des épreuves de lecture à T1 et T2

Une procédure analogue a été utilisée pour identifier les élèves en difficulté en lecture en 1ère année et en 2ème année primaire (tableau 64).

Tableau 64 : Moyennes, écart-types et centiles 25 des performances aux épreuves de lecture à T2 et T3

	Moyenne	**Ecart-type**	**Centile 25**
SG IME1	42	17,8	37
SG IME2	46	7,6	40
SG Timé2	25	6,2	20
SG Compréhension	9	2,7	8

Le tableau 65 présente le pourcentage des enfants prédits en difficulté et en réussite à partir des épreuves de lecture. En 1ère année, les élèves considérés en difficulté en lecture obtiennent à l'épreuve IME1 un score global inférieur ou égal à 37, ce qui correspond à 41 élèves sur 171 (23%). En 2ème année, 36 élèves sont considérés en difficulté à l'épreuve IME2, soit 21%. Ces élèves ont eu un score inférieur ou égal à 40 à cette épreuve. Ceux présentant des difficultés au Timé2 obtiennent des performances inférieures ou égales à 20 (37 élèves sur 171). Les élèves ayant un score inférieur ou égal à 8 à l'épreuve de compréhension ont été jugés en difficulté (28 sur 171).

Tableau 65 : Enfants prédits en difficulté et en réussite en fonction de leurs scores aux épreuves de lecture

	Prédits en difficulté	**Prédits en réussite**
SG IME1	41/171 (23%)	130/171 (76%)
SG IME2	36/171 (21%)	135/171 (78%)
SG Timé2	37/171 (21%)	134/171 (78%)
SG Compréhension	32/171 (18%)	139/171 (81%)

2. Valeur prédictive individuelle entre la classe préparatoire et la 1ère année

La méthode utilisée peut être comparée à celle de l'évaluation d'un test de dépistage. La valeur de dépistage ou de diagnostic d'un test fait appel aux notions de sensibilité et de spécificité qui sont des paramètres statistiques permettant d'apprécier quantitativement la valeur du test. Le référentiel de cette évaluation (risque positif ou négatif) est constitué par les scores des élèves aux indicateurs globaux de CP et de 1ère année tandis que le « dépistage » est défini par les performances des élèves aux différentes épreuves de lecture. Les concordances ou discordances entre les réponses fournissent des indices de sensibilité, de spécificité et de valeurs prédictives positive et négative (*cf.* tableau 66). Pour se faire, nous avons utilisé un outil de calculs médico-statistiques permettant l'évaluation de la valeur diagnostique d'une méthode de dépistage[89].

Tableau 66 : Calcul des indices et valeurs prédictives

	Difficultés en lecture	**Pas de difficultés en lecture**	
Difficultés au dépistage en CP	A (vrais positifs)	B (faux positifs)	A+B
Pas de difficultés au dépistage en CP	C (faux négatifs)	D (vrais négatifs)	C+D
	A+C	B+D	

Les tests statistiques utilisés ont permis de calculer la sensibilité, la spécificité, la valeur prédictive positive (VPP), la valeur prédictive négative (VPN) des épreuves :

- La sensibilité (A/A+C) est déterminée par le pourcentage de sujets prédits en difficulté à T1 par rapport au nombre de sujets réellement en difficulté à T+n.
- La spécificité (D/B+D) désigne le pourcentage de sujets prédits sans difficulté à T1 parmi les sujets ne présentant pas de difficulté à T+n.

[89] Cet outil statistique est disponible à l'adresse suivante : http://www.aly-abbara.com/utilitaires/statistiques/sensibilite_specificite_vpp_vpn.html.

- La valeur prédictive positive (A/A+B) renvoie au pourcentage de sujets réellement en difficulté à T+n parmi ceux prédits en difficulté à l'évaluation initiale.
- La valeur prédictive négative (D/C+D) désigne le pourcentage de sujets ne présentant réellement pas de difficulté à T+n parmi ceux prédits sans difficulté à l'évaluation initiale (T1).

Le tableau 67 indique la valeur prédictive individuelle des différents prédicteurs de CP et de 1ère année et des évaluations des enseignants sur le niveau de réussite et de difficulté des élèves à l'épreuve d'identification de mots écrits en 1ère année (IME1). Dans l'ensemble, le pourcentage de prédictions correctes au niveau des individus est élevé (entre 78% et 87%). Toutefois, des différences entre sensibilité et spécificité sont observées pour chacune des dimensions de CP. Ainsi, la sensibilité, c'est-à-dire la proportion des enfants en difficulté de lecture correctement détectés varie entre 43% et 60%. En revanche, la spécificité, c'est-à-dire la proportion de normolecteurs correctement catégorisés est plus robuste (entre 79% et 85% de classifications correctes).

Tableau 67 : Valeur prédictive individuelle des scores langagier et cognitif et du questionnaire enseignants à T1 sur les performances à l'épreuve IME1

	Sensibilité	**Spécificité**	**VPP[90]**	**VPN[91]**	**X^2**
SG langage T1	22/43	78/98	22/42	78/99	13,52
	51%	79%	52%	78%	p<.001
SG cognitif T1	18/41	109/137	18/46	109/132	9,07
	43%	79%	39%	82%	p<.01
SG Questionnaire T1	24/40	111/130	24/43	111/127	33,34
	60%	85%	55%	87%	p<.001

Les meilleures prédictions individuelles sont obtenues à partir de l'évaluation des enseignants avec une sensibilité de 60% et une valeur prédictive positive de 55%. Le niveau langagier à T1 prédit bien les difficultés de lecture puisque 51% des élèves supposés en difficulté d'après le SG langage le sont effectivement une année plus tard en lecture (sensibilité) et 52% des enfants en difficulté sont repérés par ce score (valeur prédictive positive). Les épreuves cognitives présentent la valeur prédictive la plus faible avec une sensibilité de 43% et une spécificité de 79%. Nous constatons que 39% des enfants qui ont des faibles performances

[90] Valeur prédictive positive
[91] Valeur prédictive négative

cognitives en CP vont devenir faibles lecteurs (VPP) et 82% de ceux qui ne le sont pas seront bons lecteurs (VPN).

En général, les résultats montrent que les évaluations de CP parviennent à mieux prédire la réussite ultérieure en lecture que les difficultés.

3. Valeur prédictive individuelle entre la classe préparatoire et la 2ème année

A la fin de la 2ème année primaire, l'évaluation de la lecture était basée sur deux épreuves d'identification de mots écrits : IME2 et Timé2 et une épreuve de lecture-compréhension.

3.1. Prédiction individuelle à partir de l'épreuve de lecture IME2

Les résultats présentés dans le tableau 68 indiquent que les scores globaux « cognitif » et « langage » et les évaluations des enseignants à T1 permettent de prédire de façon significative la réussite individuelle des élèves à l'épreuve d'identification de mots (IME2) deux années plus tard (spécificité et valeur prédictive négative entre 78% et 88%). En revanche, les prédictions des difficultés sont beaucoup moins fiables.

Tableau 68 : Valeur prédictive individuelle des scores langagier et cognitif et du questionnaire enseignants sur les performances à l'épreuve IME2à T3

	Sensibilité	**Spécificité**	**VPP**	**VPN**	**X^2**
SG langage T1	20/36	113/136	20/43	113/126	30,68
	55%	83%	46%	87%	p<.001
SG cognitif T1	19/35	99/126	19/46	99/115	14,49
	54%	78%	41%	86%	p<.001
SG langage T2	18/36	111/136	18/43	111/128	15,18
	50%	81%	41%	86%	p<.001
SG Questionnaire T1	22/36	112/133	22/43	112/126	22,67
	61%	84%	51%	88%	p<.001

Les meilleures prédictions sont obtenues avec le questionnaire renseigné par les enseignants (sensibilité de 61% et valeur prédictive positive de 51%). Vient ensuite le score langage à T1 qui permet de repérer 46% des enfants en difficulté. Ainsi, 46% des enfants ayant en CP un

retard au niveau des compétences langagières seront en difficulté (VPP) et 87% qui n'ont pas ce facteur de risque seront en réussite (VPN).

3.2. Prédiction individuelle à partir de l'épreuve de lecture Timé2

Lorsque le niveau de réussite ou de difficulté en lecture en 2ème année est déterminé à partir du Timé2, les prédictions concernant les futurs élèves en difficulté sont plus élevées que les précédentes (voir tableau 69) avec des sensibilités et des valeurs prédictives positives comprises entre 51% et 63%.

Tableau 69 : Valeur prédictive individuelle des des scores langagier et cognitif et du questionnaire enseignants sur les performances à l'épreuve Timé2 à T3

	Sensibilité	**Spécificité**	**VPP**	**VPN**	**X^2**
SG langage T1	21/38	107/129	21/43	107/124	22,41
	55%	82%	48%	86%	p<.001
SG cognitif T1	21/41	102/127	21/46	102/122	15,5
	51%	80%	45%	83%	p<.001
SG langage T2	22/37	92/113	22/43	92/107	22,77
	59%	81%	51%	85%	p<.001
SG Questionnaire T1	24/38	132/151	24/43	132/146	44,18
	63%	87%	55%	90%	p<.001

Le questionnaire renseigné par les enseignants permet les meilleures prédictions individuelles puisque la sensibilité et la VPP sont respectivement de 63% et de 55%. On note également la bonne sensibilité et la bonne VPP des épreuves langagières à T2 puisqu'elles repèrent 59% des enfants en difficulté. Le score global langage à T1 permet de repérer 55 % des enfants qui ont ce risque en CP et 85 % de ceux qui ne les ont pas. Les plus faibles prédictions sont obtenues avec les épreuves cognitives (sensibilités de 51% et VPP de 45%).

3.3. Prédiction individuelle à partir de l'épreuve de lecture-compréhension

Les résultats obtenus à partir de l'épreuve de compréhension confirment les données précédentes (voir tableau 70). Les meilleures prédictions sont obtenues avec le questionnaire renseigné par les enseignants et les scores globaux langage à T1 et T2. Avec ces épreuves, on

constate que plus de 60% des faibles compreneurs de 2ème année présentaient ce risque en CP (sensibilté) et plus de 80% des lecteurs en réussite ne le présentaient pas deux ans auparavant (spécificité).

Tableau 70 : Valeur prédictive individuelle des scores langagier et cognitif et du questionnaire enseignants sur les performances en lecture-compréhension à T3

	Sensibilité	**Spécificité**	**VPP**	**VPN**	**X^2**
SG langage T1	19/30	116/140	19/43	116/127	27,9
	63%	82%	44%	91%	p<.001
SG cognitif T1	14/33	104/136	14/46	104/123	4,79
	42%	76%	30%	84%	p<.05
SG langage T2	19/31	109/133	19/43	109/121	24,3
	61%	81%	44%	90%	p<.001
SG Questionnaire T1	21/33	115/137	21/43	115/127	31,86
	63%	83%	48%	90%	p<.001

4. Prédictions différentielles

Dans les analyses précédentes, nous avons cherché à étudier la valeur prédictive individuelle des évaluations cognitives et langagières et des évaluations des enseignants en CP sur le niveau de lecture en 1ère année et en 2ème année. Notre objectif ici est de procéder à une analyse différentielle régressive afin de savoir si les différences individuelles observées en 2ème année dans les épreuves de lecture correspondaient en CP à des profils spécifiques. Autrement dit, les enfants identifiés en difficulté dans l'apprentissage de la lecture en 2ème année éprouveraient-ils dès la CP des difficultés dans les épreuves auxquelles nous les avions soumis ? Ceci confirmerait l'effet Matthieu mis en évidence par Stanovich (1986) : les enfants qui, dès le début de l'apprentissage de la lecture, éprouvent des difficultés ne peuvent pas bénéficier comme les autres enfants de l'enrichissement induit par le contact avec l'écrit. Par conséquent, les riches s'enrichissent encore plus tandis que les pauvres s'appauvrissent.

Les enfants en 2ème année primaire ont été regroupés en trois groupes : faibles lecteurs (FL), lecteurs moyens (ML) et bons lecteurs (BL), selon leurs résultats aux épreuves d'identification de mots (IME2 et Timé2). Les enfants dits faibles ont des scores au-dessous du centile 25, ceux du groupe moyen ont des performances comprises entre le centile 25 et le

centile 75 et le groupe de bons lecteurs a des performances supérieures au centile 75. Ainsi, les FL ont un score inférieur à 40 à l'épreuve IME2 et à 20 à l'épreuve Timé2 (N=44). Les ML ont un score compris entre 40 et 53 à l'épreuve IME2 et entre 20 et 32 à l'épreuve Timé2 (N=74). Les BL ont un score supérieur à 53 à l'épreuve IME2 et à 32 au Timé2 (N=53). Nous essayerons ainsi de remonter dans le temps et de voir si les trois groupes se différenciaient déjà en CP sur les domaines en relation avec l'apprentissage de la lecture.

4.1. Comparaisons effectuées en classe préparatoire

Nous présentons dans la figure 14 les résultats obtenus par les trois groupes de lecteurs de 2ème année aux épreuves langagières et au questionnaire renseigné par les enseignants en fin de CP.

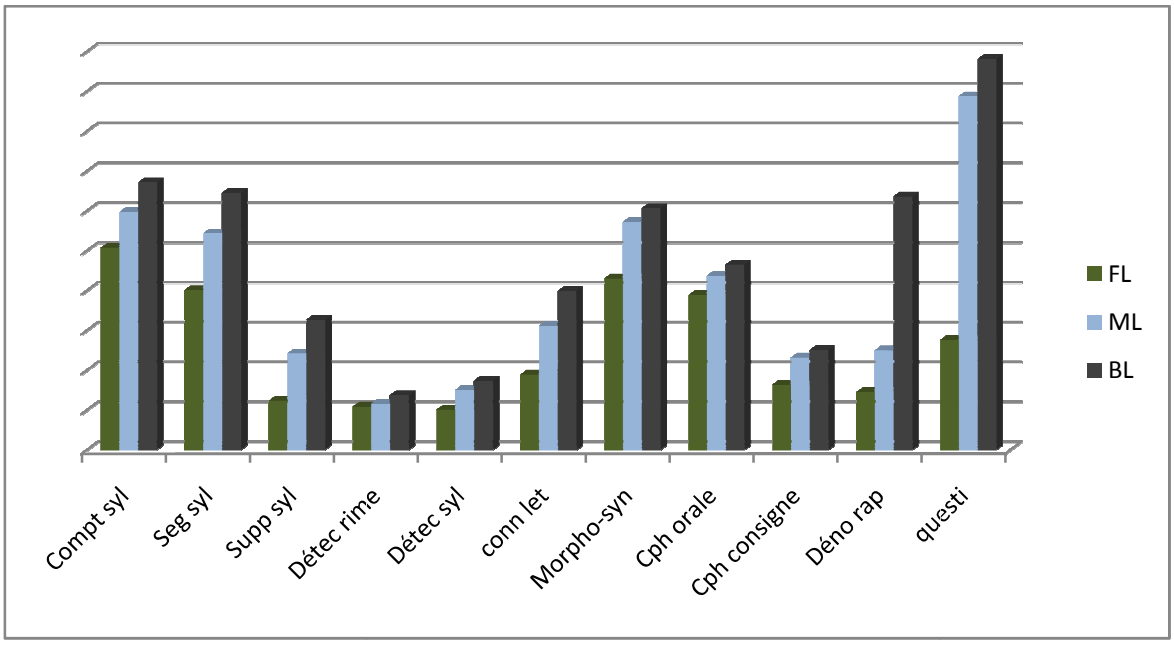


Figure 14 : Performances obtenues par les trois groupes de lecteurs aux différentes épreuves langagières et au questionnaire enseignants en CP

De façon générale, nous constatons que les BL ont dans la plupart des épreuves langagières des performances supérieures à celles des ML et FL. D'autre part, les FL ont des scores inférieurs aux lecteurs moyens, même si l'on peut observer quelques cas où leurs performances sont équivalentes. Pour chaque épreuve, des analyses de variance (ANOVA) sur les performances obtenues par les trois groupes ont été effectuées. Les résultats de ces comparaisons figurent dans le tableau 71. Les différences sur les épreuves langagières sont globalement significatives à l'analyse de variance et permettent de distinguer

significativement les trois groupes. Seules deux épreuves ne permettent pas de discriminer les trois groupes de lecteurs : la rapidité de dénomination et la détection de l'intrus de rime.

Tableau 71: Différences significatives sur les compétences langagières à T1 selon les niveaux de performances aux épreuves d'identification de mots à T2 (faibles=FL ; moyens=ML ; bons=BL)

	FL/BL	**FL/ML**	**ML/BL**	**F**
Déno Rapide				2,48 ns
Compt syl	*	*		8,29 P<.001
Seg syl	*	*	*	11,39 P<.001
Supp syl	*	*	*	14,55 P<.001
Détec Rime				2,58 ns
Détec Syl	*	*		13,99 P<.001
Morpho-syn	*	*		14,10 P<.001
Conn let	*	*	*	21,86 P<.001
Cph orale	*	*		5,35 P<.01
Cph consignes	*	*	*	36,25 P<.001
Questionnaire	*	*		31,46 P<.001

* s à p<.05

Lorsqu'on procède à des comparaisons deux à deux, nous constatons que les différences les plus nombreuses et les plus significatives concernent plutôt les groupes extrêmes (BL, FL). Des différences très sensibles sont également observées entre les faibles lecteurs (FL) et les lecteurs moyens (ML). Il est difficile cependant de distinguer statistiquement les lecteurs moyens (ML) et les bons lecteurs (BL) sur l'ensemble des épreuves où des différences significatives apparaissent globalement entre les trois groupes. Les différences entre moyens et bons ne concernent qu'un nombre limité de compétences : segmentation et suppression syllabiques, connaissance des lettres et compréhension de consignes.

Le questionnaire renseigné par les enseignants permet de différencier les trois groupes. Les enfants classés comme bons lecteurs sont évalués plus positivement que les lecteurs moyens et les faibles lecteurs. En revanche, les ML se différencient très peu en CP des BL et font l'objet de la part des enseignants d'évaluations identiques. De manière générale, c'est sur les

épreuves les plus difficiles et les plus prédictives de la réussite en lecture que les différences se manifestent entre les trois groupes et ceci, dès la classe préparatoire.

La figure 15 présente les performances observées pour les trois groupes de lecteurs à chacune des épreuves cognitives. L'analyse globale montre que toutes les épreuves cognitives permettent de discriminer les trois groupes. Nous avons cependant effectué sur chaque épreuve une analyse de variance et une comparaison deux à deux entre les groupes.

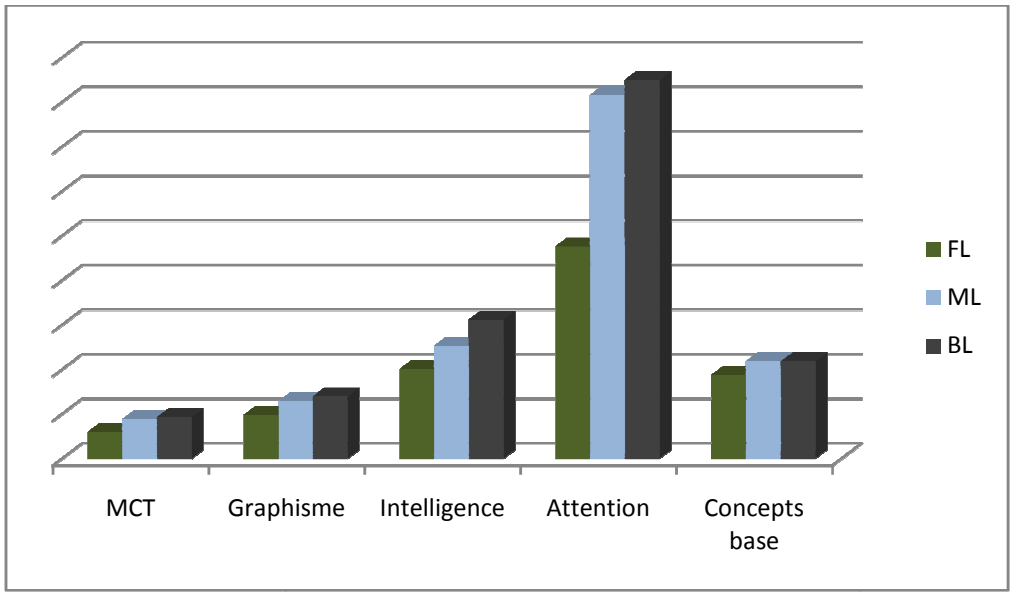


Figure 15 : Performances obtenues par les trois groupes de lecteurs de 2ème année aux différentes épreuves cognitives en CP

Les résultats présentés dans le tableau 72 révèlent des différences importantes et significatives entre les BL et les FL d'une part et entre les FL et les ML d'autre part dans toutes les épreuves cognitives. Toutefois, comme dans les évaluations précédentes, il est difficile de distinguer statistiquement les lecteurs moyens et les bons lecteurs. Seule l'épreuve d'intelligence non verbale permet de discriminer les deux groupes de lecteurs.

Tableau 72 : Différences significatives sur les compétences cognitives à T1 selon les niveaux de performances aux épreuves d'identification de mots à T3 (faibles=FL ; moyens=ML ; bons=BL)

	FL/BL	FL/ML	ML/BL	F
Int	*	*	*	16 P<.001
Concepts de base	*	*		14,99 P<.001
MCT	*	*		12,75 P<.001
Graph	*	*		13,87 P<.001
Bar1	*	*		13,99 P<.001
Bar2	*	*		5,42 P<.01

* s à p<.05

4.2. Comparaisons effectuées en 1ère année

Nous présentons dans la figure 16 les résultats obtenus par les trois groupes de lecteurs sur chacune des épreuves passées en 1ère année primaire (IME1, détection du phonème initial et final et jugement des liens morphologiques).

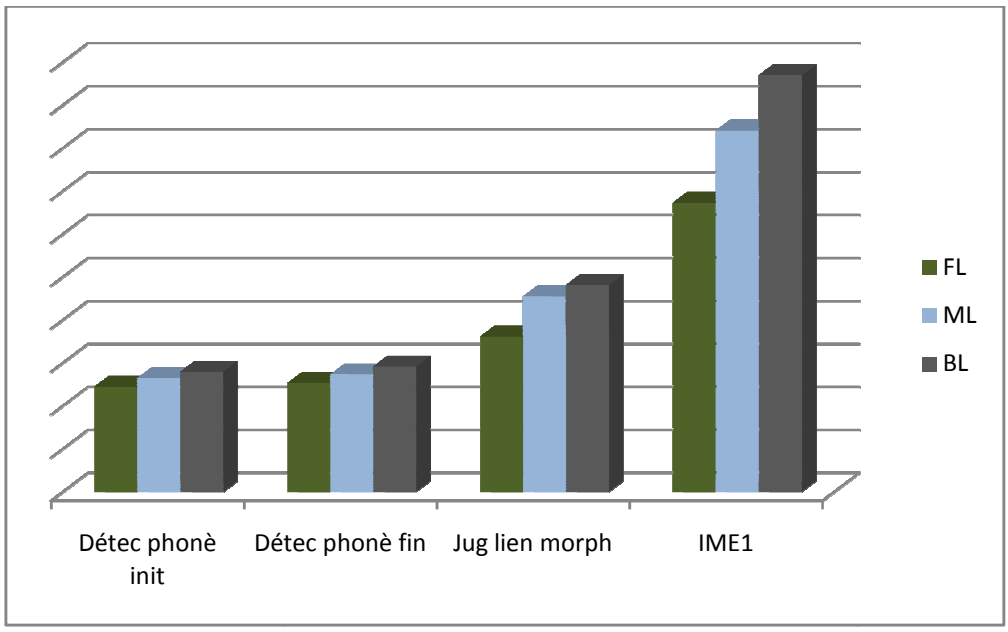


Figure 16 : Performances obtenues par les trois groupes de lecteurs aux différentes épreuves passées en 1ère année

Comme nous avons pu le constater antérieurement, les BL ont dans toutes les épreuves des performances supérieures à celle des ML et des FL. La comparaison des trois groupes à l'aide des analyses de variance a permis d'observer des différences significatives à p<.001 (voir tableau 73). Les trois groupes diffèrent significativement entre eux sur les compétences en identification de mots écrits en 1ère année (IME1). Par conséquent, on observe une certaine stabilité dans la différenciation des performances en lecture des trois groupes de lecteurs. La détection de l'intrus phonémique et le jugement de liens morphologiques différencient également les trois groupes. Les relations vont toujours dans le même sens : les bons lecteurs ont des scores significativement plus élevés que les moyens et les faibles lecteurs. Toutefois, seule l'épreuve de détection du phonème initial permet de distinguer les lecteurs moyens des bons lecteurs.

Tableau 73 : Différences significatives sur les épreuves passées à T2 selon les niveaux de performances aux épreuves d'identification de mots à T3 (faibles=FL ; moyens=ML ; bons=BL)

	FL/BL	FL/ML	ML/BL	F
IME1	*	*	*	61,01 P<.001
Phonè init	*	*	*	15,73 P<.001
Phonè fin	*	*		12,35 P<.001
Jug morph	*	*		26,71 P<.001

* s à p<.05

5. Discussion

Dans les chapitres précédents, nous avons essayé d'étudier la valeur prédictive globale des mesures langagières et cognitives et du questionnaire renseigné par les enseignants sur le futur niveau de lecture des enfants. Les résultats indiquent clairement, tout d'abord, que les variables prédictives mesurées en fin de CP et de 1ère année apportent une contribution non négligeable à la variance en lecture. Les évaluations des enseignants présentent également une bonne valeur prédictive globale des performances en lecture en 1ère année et en 2ème année quelle que soit l'épreuve de lecture utilisée. Les évaluations prédictives globales restent cependant insuffisantes lorsqu'il s'agit de prédire le devenir des élèves pris individuellement.

Dans ce chapitre, nous avons essayé de mesurer la valeur prédictive individuelle des différentes mesures dans une perspective de prévention. En d'autres termes, nous avons essayé de déterminer la proportion des enfants en difficulté de lecture correctement détectés par nos mesures prédictives (sensibilité). Le repérage des enfants risquant de se trouver en difficulté dans l'apprentissage de la lecture constitue une étape nécessaire à une prise en charge précoce et efficace des difficultés scolaires.

Les résultats indiquent que le pourcentage de prédictions correctes au niveau des individus est élevé. Les prédicteurs sélectionnés pour cette analyse permettent à eux seuls de classer correctement plus de 80% des enfants en fonction de leur niveau de lecture en fin de 1ère année et de 2ème année primaire. Toutefois, la prédiction est moins fiable pour les enfants qui deviennent faibles lecteurs que pour ceux qui sont jugés comme furturs bons lecteurs. Les sensibilités et les valeurs prédictives positives (VPP) varient entre 39% et 55% en 1ère année et entre 41% et 55% en 2ème année. En revanche, les spécificités et les valeurs prédictives négatives (VPN) sont plus robustes : entre 78% et 87% de classifications correctes en 1ère année et entre 83% et 90% en 2ème année. Ces résultats sont en concordance avec ceux obtenus par Peterson & Elbro (1999) avec des enfants danois et ceux obtenus par Piquard-Kipffer (2003) et Guimard & Florin (2007) avec des enfants français. L'étude de Zorman & al. (2004), réalisée entre la grande section de maternelle et le CE1, montre la faible capacité des évaluations de la conscience phonologique et des compétences en langage oral à repérer dès la maternelle les futurs enfants en difficulté (sensibilité et VPP entre 30% et 60%). Par contre, la spécificité et la VPN de ces évaluations sont très élevées (entre 80% et 92%).

Les meilleures prédictions individuelles du niveau de lecture à T2 sont obtenues à partir de l'évaluation des enseignants avec une sensibilité de 60% et une VPP de 55% et le niveau langagier qui permet de repérer plus de 50% des enfants en difficulté. Toutefois, le score cognitif ne semble pas avoir une bonne capacité prédictive individuelle puisque la sensibilité et la VPP se situent au dessous de 50% (respectivement 43% et 39%). Ces données semblent indiquer que les évaluations des comportements scolaires effectuées par les enseignants et les évaluations langagières en fin de CP prédisent mieux le risque d'échec en 2ème année que les évaluations cognitives réalisées à la même période. Ces résultats sont globalement conformes à ceux de Guimard & Florin (2007) qui ont trouvé que les meilleures prédictions concernant les futurs élèves en difficulté sont obtenues à partir de l'indicateur « langage » de la BREV avec une sensibilité et une VPP de 56% et 50% et le score globale au questionnaire renseigné par les enseignants (sensibilité de 44% et VPP de 39%). Les plus faibles prédictions

individuelles proviennent du score cognitif de la BREV avec une sensibilité de 31% et une VPP de 29,5%.

Nous avons pu constater également que la qualité de la prédiction ne varie pas en fonction de l'épreuve de lecture choisie en variable dépendante, ni en fonction du niveau scolaire. Ainsi, les prédictions individuelles obtenues avec les différentes épreuves de lecture en 1ère année et en 2ème année (IME1, IME2, Timé2 et compréhension) sont presque équivalentes. Cependant, il faut noter que les élèves considérés en difficulté ou en réussite par les enseignants en CP ne sont pas nécessairement les mêmes d'après les scores langagiers et cognitifs. Il en va de même en 1ère année et 2ème année lorsqu'on observe la répartition des deux groupes d'élèves selon les différentes épreuves de lecture.

Nous avons essayé dans ce chapitre d'analyser la valeur différentielle des différentes mesures afin d'envisager la construction d'instruments de dépistage des difficultés en lecture. Il était possible ainsi de repérer en fin de 2ème année les enfants bons lecteurs (BL), lecteurs moyens (ML) et faibles lecteurs (FL), puis par une analyse différentielle rétrospective, de remonter dans le temps pour voir si ces trois groupes se différenciaient déjà en classe préparatoire et en 1ère année sur les épreuves auxquelles nous les avions soumis. De manière générale, nous constatons que les évaluations cognitives et langagières de CP permettent assez nettement de prédire les contrastes entre élèves BL et FL que l'on peut observer à l'école primaire. Les BL ont, dans toutes les épreuves des performances supérieures à celles des ML et des FL. D'autre part, dans de nombreux cas, les FL ont des performances inférieures à celles des ML. Il est difficile cependant de distinguer les ML et les BL. Certaines épreuves langagières ont une incidence visible sur le développement différentiel des capacités en lecture. Seules les épreuves de dénomination rapide et de détection de l'intrus de rime ne permettent pas de différencier les trois groupes de lecteurs. Les épreuves à propos desquelles une différence apparait sont celles dont nous avons pu constater la liaison avec les épreuves de lecture en 1ère année et en 2ème année et qui se trouvent les plus prédictives de la réussite en lecture (connaissances phonologiques et morphologiques, connaissance des lettres, etc.). Ces résultats sont conformes à ceux trouvés par Lecoq (1991) qui a effectué une analyse rétrospective pour voir comment les performances des trois groupes de lecteurs (BL, ML et FL) ont évolué depuis la maternelle aussi bien dans les épreuves phonologiques, mnésiques et de rapidité de dénomination que lors des tests de lecture effectués en CP. Les données montrent que, de manière générale, c'est sur les épreuves les plus difficiles et les plus prédictives de la réussite en lecture, que les différences se manifestent entre les trois groupes et ceci dès la maternelle.

Ainsi, les bons lecteurs ont, dans toutes les épreuves phonologiques, une performance supérieure à celle des lecteurs moyens et des faibles lecteurs. Les résultats obtenus sur les épreuves de mémoire sont beaucoup moins probants : seules les tâches les plus souvent corrélées avec les épreuves phonologiques permettent de différencier les trois groupes de lecteurs. En ce qui concerne les épreuves de rapidité, l'auteur a constaté que la vitesse d'articulation n'avait pas d'incidences visibles sur le développement différentiel des capacités en lecture. Nous avons pu constater également que les évaluations des enseignants permettent de prédire les contrastes entre BL et FL. En revanche, les différences entre ML et BL paraissent plus difficiles à établir. Ces résultats vont dans le sens de ceux établis par Guimard & Florin (2001) qui ont étudié la valeur différentielle des évaluations comportementales des enseignants. Les analyses rétrospectives indiquent que les élèves de CP[92] faibles en lecture et les élèves de CE2[93] faibles en lecture-compréhension et aux évaluations nationales de mathématiques, faisaient l'objet en moyenne section de maternelle d'évaluations comportementales plus négatives que les élèves moyens et forts.

Cette analyse différentielle se révèle très intéressante puisqu'elle vient confirmer les analyses de corrélation et de régressions multiples qui ont été menées jusqu'à présent et met en évidence que ce sont les capacités les plus liées à l'apprentissage de la lecture qui permettent de différencier significativement les trois groupes de lecteurs. Les différences relevées entre les trois groupes en classe préparatoire se maintiennent en 1ère année sur l'épreuve de lecture et les épreuves de détection de phonèmes et de jugement de liens morphologiques. Comme les BL et les ML réussissent à progresser de la CP à la 2ème année, les FL quant à eux persistent dans leurs difficultés, ce qui renvoie à l'effet Matthieu : le bon lecteur deviendra encore meilleur, le faible lecteur encore plus faible (Stanovich, 1986). Il faudrait toutefois disposer d'informations relevées d'un niveau scolaire supérieur pour conclure avec assurance que les difficultés des enfants FL se maintiendraient dans le temps.

Cette prédiction différentielle suggère qu'il est possible d'identifier dès la CP des enfants présentant des profils « à risque » pour les apprentissages ultérieurs. Le repérage des enfants risquant de se trouver en difficulté dans les apprentissages fondamentaux et l'identification de leurs difficultés spécifiques en fin de CP constituent les deux étapes nécessaires à une prise en charge précoce et efficace des difficultés en lecture.

[92] Correspond à la 1ère année dans le système éducatif tunisien.
[93] Correspond à la 3ème année dans le système éducatif tunisien.

Conclusion générale

L'apprentissage réussi de la lecture est un processus complexe qui constitue l'un des défis majeurs des premières années de la scolarisation. Reconnus par l'OMS (Organisation Mondiale de la Santé), les troubles du langage oral et écrit et les troubles d'apprentissage nécessitent des mesures de prévention qui allient repérage des enfants à risques, dépistage des difficultés et évaluation précise permettant de proposer éventuellement les aides nécessaires et adaptées.

En Tunisie, on observe un intérêt particulier pour ce domaine de recherche étant donné la prise de conscience générale de la nécessité de repérer les enfants à risque dès la classe préparatoire dans une perspective de prise en charge précoce et adaptée. La connaissance des mécanismes impliqués dans le fonctionnement cognitif et langagier permettra de créer des outils d'évaluation indispensables au repérage et à la prédiction/prévention des difficultés des enfants en contexte scolaire. Notre étude s'inscrit dans cette démarche : déterminer les compétences précoces impliquées dans la maîtrise de l'écrit et permettre précocement le repérage d'enfants à risque pour leur proposer les aides nécessaires.

L'objectif de cette étude longitudinale était d'examiner la valeur prédictive globale et individuelle des compétences langagières et cognitives précoces et des évaluations des enseignants sur les performances ultérieures en lecture. Elle visait précisément à déterminer les meilleurs prédicteurs de la réussite ou de l'échec dans l'apprentissage de la lecture en langue arabe. Nous rappelons à ce stade les principales questions soulevées par notre problématique de départ:

Quelles sont les meilleurs prédicteurs de l'apprentissage de la lecture en langue arabe ?

Dans quelle mesure les compétences langagières et cognitives évaluées en classe préparatoire peuvent-elles contribuer à prédire l'échec ou la réussite en lecture ?

Est-il possible de modéliser les relations prédictives entre les compétences précoces et les performances ultérieures en lecture ?

Dans quelle mesure l'échec et la réussite en lecture peuvent-ils être prédits par les enseignants en classe préparatoire ? Sur quels indicateurs peuvent-ils fonder leurs évaluations ?

Les mesures utilisées dans cette recherche présentent-elles une bonne valeur prédictive individuelle ? Dans quelle mesure contribuent-elles au dépistage des difficultés de l'apprentissage de la lecture ?

Ce sont des questions auxquelles nous avons tenté de répondre en nous appuyant sur l'étude d'une cohorte de 171 élèves tunisiens suivis de la classe préparatoire[94] à la 2ème année primaire. En classe préparatoire, différentes épreuves langagières (compétences phonologiques, compétences morphologiques, connaissance des lettres, compréhension orale, compréhension de consignes, dénomination rapide) et cognitives (intelligence, mémoire, graphisme, attention, concepts de base) ont été administrées aux élèves. Les évaluations des comportements et des compétences scolaires des élèves ont été réalisées grâce à un questionnaire renseigné par les enseignants de CP. L'efficience en lecture a été évaluée chez les mêmes enfants, en 1ère année et en 2ème année, par des épreuves d'identification de mots écrits et une épreuve de lecture-compréhension.

Les meilleurs prédicteurs de l'apprentissage de la lecture en langue arabe

Un premier objectif de cette recherche était d'étudier le poids prédictif des compétences langagières et cognitives précoces sur les performances ultérieures en lecture. De nombreux travaux de recherche ont montré que cet apprentissage est favorisé par le développement d'habiletés langagières orales ainsi que par l'émergence de capacités métalinguistiques (Demont & Gombert, 2007 ; Ecalle & Magnan, 2007b ; Gombert, 2003 ; 1990 ; Gombert & al., 2004 ; Storch & Witehurst, 2002 ; Sprenger-Charolles & Casalis, 1996). Deux autres prédicteurs importants ressortaient également, la connaissance des lettres et la dénomination rapide (Castel & al., 2008 ; Lecoq, 1991 ; Schatschneider & al., 2004). Toutefois, l'importance accordée aux connaissances phonologiques dans l'apprentissage de la lecture a pour conséquence de négliger l'influence d'autres types de compétences qui ne sont pas prises en compte par les chercheurs. En effet, des études de plus en plus nombreuses montrent que dès les premiers stades d'apprentissage de la lecture, les enfants sont sensibles à la structure morphologique des mots lus et sont capables de la manipuler (Abu-Rabia, 2007 ; Besse, 2007 ; Besse & al., 2007 ; Boukadida, 2008 ; Marec-Breton, 2009 ; 2003). Les recherches menées sur la langue arabe apportent des résultats congruents avec l'hypothèse d'une intervention précoce et privilégiée de la morphologie lors de la lecture débutante (Boukadida, 2008 ; Boukadida & al., 2009).

[94] La classe préparatoire en tunisie correspond à la grande section de maternelle (GSM) en France.

La revue de la littérature a révélé en effet que les compétences langagières constituent des prédicteurs très puissants des capacités de lecture ultérieures. Kirby, Desrochers, Roth & Lai (2008) ont passé en revue la littérature scientifique sur les facteurs les plus prédictifs de l'apprentissage de la lecture. Ils dégagent ainsi six construits qui s'avèrent les plus corrélés à l'efficience ultérieure en lecture : la conscience phonologique, la dénomination rapide, les processus orthographiques, la conscience morphologique et le niveau de voacbulaire. Dans la plupart des cas, la corrélation demeure significative lorsque les autres facteurs prédictifs sont contrôlés. La majorité des données disponibles concerne la langue anglaise et rares sont les études qui ont tenté d'étudier les variations interlangues. Nous avons essayé dans cette recherche d'apprécier ce qui dépend des caractéristiques de la langue arabe et de voir si les prédicteurs devraient différer en fonction du système linguistique de chaque langue.

Les résultats des analyses descriptives montrent que, dès la classe préparatoire, les enfants tunisiens sont amenés à découvrir certaines règles de fonctionnement du code écrit et à en prendre conscience. Les capacités à identifier et à manipuler les syllabes apparaîssent être acquises de façon précoce et ne posent pas de problèmes particuliers aux enfants dès le début de l'apprentissage de la lecture. Les enfants prélecteurs sont également en mesure de reconnaître les lettres de l'alphabet et maîtrisent les règles flexionnelles des phrases arabes. En revanche, il faut attendre la 1ère année de l'apprentissage formel de la lecture pour qu'ils soient en mesure de manipuler les unités phonémiques et d'intégrer la notion d'appartenance à une famille morphologique. L'entrée dans l'écrit fournit aux enfants des indices importants pour la compréhension de la morphologie de leur langue.

Les résultats obtenus par des analyses corrélationnelles vont tout à fait dans le sens des connaissances théoriques actuelles sur les compétences qui sous-tendent l'apprentissage de la lecture. La reconnaissance des mots écrits mobilise des connaissances d'ordre phonologique et morphologique et des connaissances du nom des lettres. Pour apprendre à lire, l'enfant doit avoir également des compétences mnésiques et graphiques et une maîtrise des concepts de base qui, par ailleurs, continuent de se développer grâce à l'apprentissage de la lecture. La rime, qui apparaît comme une unité privilégiée en anglais ne semble effectivement pas bénéficier du même statut en arabe car l'analogie attaque-rime entre les mots n'est pas pertinente en arabe. Ceci est vérifié par les faibles corrélations entre l'épreuve de détection de l'intrus de rime et les épreuves de lecture quel que soit le niveau scolaire. En effet, l'identification de mots en arabe est basée sur le squelette consonantique et nécessite une décomposition morphologique en racine-schème et non pas une segmentation attaque-rime

(Ammar, 1997). Ces résultats confirment ceux obtenus par Boukadida (2008) qui montrent que l'utilisation des connaissances sur les rimes lors de la lecture chez l'apprenti lecteur arabophone n'est pas pertinente au début de l'apprentissage de la lecture. De même, Besse & al. (2007) ont trouvé que les scores à l'épreuve implicite sur la rime s'avèrent inférieurs à ceux obtenus aux deux épreuves de préférence sur le lien morphologique chez les arabophones mais pas chez les lusophones.

Parmi l'ensemble des compétences évaluées en CP, certaines apparaissent comme particulièrement prédictives du niveau global d'acquisition des élèves en lecture. Il semberait, conformément à notre hypothèse, que les variables les plus prédictives des performances en lecture appartiennent au langage oral et écrit. Au niveau de la 1ère année, les compétences les plus prédictives des performances en lecture et qui expliquent une part propre de la variance des performances en identification de mots se regroupent dans trois grands domaines : connaissance des lettres, conscience phonologique et morpho-syntaxe. Ces résultats confirment les conclusions de nombreuses études qui ont montré l'importance de la connaissance des lettres et des compétences phonologiques en tant que prédicteurs puissants du développement ultérieur des compétences en lecture (Demont & Gombert, 1996 ; Ecalle & al., 2009 ; Lecoq, 1991 ; Muter & al., 2004 ; 1998 ; Schatschneider & al., 2004 ; Storch & Withehurst, 2002). Notre étude montre que les connaissances précoces sur la morphologie flexionnelle pourraient constituer un bon prédicteur de la réussite en identification de mots en langue arabe. Les travaux portant sur les liens entre conscience morphologique et apprentissage de la lecture se sont intéressés essentiellement au domaine de la morphologie dérivationnelle parce qu'elle intègre différentes dimensions linguistiques (phonologique, sémantique et syntaxique) (Boukadida, 2008 ; Carlisle, 1995 ; Carlisle & Nomanbhyne, 1993). Différentes études comme celles de Lecoq & al. (1996) indiquent que les connaissances morpho-syntaxiques affectent essentiellement les performances aux tâches de compréhension.

Le développement des capacités d'identification de mots et de compréhension en 2ème année est prédit également par différentes compétences langagières. En particulier, la sensibilité phonologique et la connaissance des lettres sont des prédicteurs puissants des variations des compétences d'identification de mots. Toutefois, les connaissances morpho-syntaxiques ne sont pas aussi importantes pour l'identification de mots mais semblent prédire de façon significative la compréhension (15% de la variance expliquée). Nous pouvons penser qu'à partir du moment où l'habileté de décodage est suffisamment développée pour permettre une

identification correcte et rapide des mots écrits, les compétences de plus haut niveau telles que la conscience morpho-syntaxique cessent d'être fonctionnelles au niveau des processus de l'identification des mots et s'investissent dans la lecture-compréhension.

Les données montrent que la connaissance des lettres, les connaissances phonologiques et, dans une moindre mesure, la compréhension de consignes, mesurées en CP ont un assez bon pouvoir prédictif sur l'acquisition de la lecture en fin de 2ème année (2 ans plus tard). La connaissance des lettres et la segmentation syllabique en CP expliquent, à elles seules, 24% de la variance à l'épreuve IME2 et 34 % de la performance au Timé2 en fin des deux années d'apprentissage de la lecture. De même, la compréhension de consignes en CP explique une part significative de la variance du niveau de langage écrit en 2ème année. Il en résulte que cette compétence mérite d'être évaluée et travaillée en CP en vue, par exemple, de repérer au sein de la classe les difficultés de certains enfants dans ce domaine. En effet, la compréhension de consignes constitue une des conditions de de la réussite dans les différentes activités scolaires puisque la majorité d'entre elles sont introduites par des consignes.

Le niveau de lecture en 2ème année est significativement corrélé aux connaissances morphologiques dérivationnelles et aux connaissances phonémiques mesurées en 1ère année. Les performances des élèves dans la tâche de morphologie sont même plus corrélées avec les performances en lecture une année plus tard que ne le sont les performances dans la manipulation des unités phonémiques. Les analyses de régression révèlent que les performances des élèves dans la tâche de jugement des liens morphologiques expliquent une part importante de la variance en identification de mot (IME2 : 13% de variance ; Timé2 : 21% de variance) et en lecture-compréhension (11%) et apparaissent être un meilleur prédicteur que les épreuves de détection de l'intrus phonémique. Ces résultats vont dans le sens de ceux obtenus par les études menées dans différentes langues qui mettent en évidence un lien entre le développement des connaissances morpho-dérivationnelles et l'acquisition de la lecture, dès la première année d'apprentissage (Boukadida, 2008 ; Boukadida & al., 2009 ; Casalis & Louis-Alexandre, 2000 ; Marec-Breton, 2003).

En arabe, la prise en compte des aspects morphologiques est particulièrement déterminante pour apprendre à lire compte tenu des caractéristiques de son système orthographique. Une des spécificités de l'orthographe arabe est qu'elle opère une distinction nette entre les consonnes et les voyelles. En langue arabe, le mot a une structure bidimensionnelle. Il est constitué d'une racine composée seulement de consonnes alors que les voyelles sont

représentées par des signes diacritiques, non incorporés au squelette consonantique, fournissant une forme lexicale et grammaticale au mot. Ces particularités expliquent l'implication des connaissances morphologiques dérivationnelles dans l'apprentissage de la lecture en arabe.

Si, comme nous l'avons vu, les épreuves cognitives prédisent nettement moins la réussite ou l'échec en lecture, nous avons pu constater par les analyses de corrélation et l'analyse en composantes principales qu'elles pouvaient avoir une relation étroite avec certaines compétences langagières. Chacune des épreuves langagières met en jeu les compétences cognitives nécessaires pour aborder l'écrit : maîtrise des concepts de base et compétences mnésiques, notamment pour les épreuves de conscience phonologique ; des capacités grapho-motrices pour la connaissance des lettres. Nous pouvons ainsi considérer que les compétences cognitives contribuent indirectement à l'explication de la variance en lecture.

Par ailleurs, les faibles résultats obtenus à la tâche de dénomination rapide montrent que les enfants, avant le début de l'apprentissage de la lecture, trouvent des difficultés à accéder de façon rapide et précise aux représentations phonologiques dans leur lexique mental compte tenu de la situation de diglossie. En effet, nous avons constaté que la majorité des enfants ont très souvent tendance à recourir à l'arabe standard (associé à l'école et aux apprentissages scolaires), ce qui a causé une lenteur dans l'accès lexical et la récupération de la représentation phonologique du mot. En plus, l'analyse des réponses des enfants montrent que, dans la majorité des cas, les élèves tendent à donner des réponses combinant des éléments phonologiques des deux registres[95].

Enfin, un dernier point dégagé de l'ensemble des résultats est que la qualité de la prédiction peut dépendre de l'épreuve de lecture choisie en variable dépendante. Les meilleures prédictions du niveau en identification de mots en 2ème année sont obtenues avec le Timé2. Ces variations peuvent s'expliquer par la nature différente des compétences en lecture évaluées par les deux épreuves et par leur plus ou moins grande sensibilité dont rend compte la courbe de distribution des scores. Ces résultats soulignent l'intérêt de bien choisir les outils d'évaluation utilisés pour pédire les difficultés ultérieures en lecture et de s'assurer des qualités psychométriques de ces instruments. L'utilisation d'une seule épreuve de lecture peut paraître insuffisante pour identifier avec assurance les enfants en difficulté.

[95] Par exemple, l'enfant ajoute le tanwin, réservé à l'arabe standard, lorsqu'il dénomme les images en arabe dialectal (i.e., صبّاطٌ ; كرهبةٌ).

Les résultats de cette étude sont en partie conformes aux données de la littérature internationale mais soulignent qu'une évaluation prédictive doit tenir compte des caratéristiques de la langue et du contexte d'apprentissage. L'ensemble de ces résultats pourrait suggérer de développer les démarches qui contribuent à repérer rapidement et précocement les enfants à risque de difficulté dans l'apprentissage de la lecture afin de leur proposer éventuellement des aides nécessaires et adaptées.

Elaborer un modèle structurel des prédicteurs de l'apprentissage de la lecture en langue arabe

Un autre objectif poursuivi dans cette étude était de tester un modèle des prédicteurs de l'apprentissage de la lecture en langue arabe. Un modèle d'équations structurelles teste les relations linéaires hypothétiques entre plusieurs variables. Chaque relation testée doit être par conséquent fondée sur des bases théoriques. La spécification d'un modèle théorique consiste d'une part à déterminer le modèle de mesure c'est-à-dire à décrire finement une variable latente par ses indicateurs (variables observées), et d'autre part à déterminer le modèle structurel en essayant d'expliquer les relations linéaires pouvant s'établir entre plusieurs variables latentes. Les effets directs et indirects des variables latentes explicatives ou variables exogènes sur la variable latente à expliquer ou variable endogène peuvent être évalués par les coefficients de causalité qui relient les différents construits du modèle à tester. Ces techniques tiennent compte également des erreurs de mesures propres aux variables exogènes. L'évaluation du modèle consiste à apprécier la qualité de l'ajustement du modèle théorique aux données empiriques. Un bon ajustement est une condition de validité interne des résultats qui pourront ensuite être interprétés. Un modèle dont la qualité d'ajustement est médiocre ne permet pas l'acceptation et l'interprétation des estimations des relations linéaires.

Le recours à la modélisation par les équations structurelles a permis de prendre en considération les relations linéaires entre les connaissances phonologiques et morphologiques précoces et les performances ultérieures en lecture. Les recherches longitudinales ont révélé que certaines capacités initiales des enfants peuvent prédire le déroulement ultérieur de l'apprentissage de la lecture (*cf.* partie théorique ; chapitre 4) et qu'il existe une continuité développementale entre les compétences préscolaires et les apprentissages scolaires. Les relations prédictives signalent aussi que certaines habiletés préscolaires ont davantage de poids que d'autres sur l'apprentissage de l'écrit (Storch & Whitehurst, 2002). Les habiletés de sensibilité phonologique et les habiletés langagières sont à cet égard reconnues comme des

prédicteurs puissants de cet apprentissage. Toutefois, les études qui ont essayé de modéliser la relation prédictive entre les compétences langagières précoces et l'apprentissage de la lecture (Muter & al., 2004 ; Leong & al., 2005 ; Storch & Whitehurst, 2002) n'ont pas tenu compte du rôle de la dimension morphologique. L'hypothèse postulée suggérait que le niveau de connaissances phonologique et morphologique à T1 explique une part des performances en lecture en arabe.

Le modèle proposé dans cette étude considère que les performances en lecture en langue arabe sont expliquées par deux variables latentes : connaissances phonologiques et connaissances morphologiques. Plus précisément, nous avons souhaité évaluer l'effet direct des connaissances phonologiques sur le niveau de lecture et son effet indirect, par la médiation des connaissances phonémiques à T2. La décision d'intégrer la variable « connaissances morphologiques » dans le modèle est basée sur les résultats de plusieurs recherches qui ont mis en évidence la relation étroite entre la dimension morphologique et l'apprentissage de la lecture en arabe (Ammar, 1997 ; Boukadida, 2008 ; Boukadida & al., 2009 ; Besse & al., 2007 ; Belajouza, 2003).

Les résultats de l'analyse utilisant la modélisation par les équations structurelles supportent l'hypothèse d'une liaison positive entre les connaissances phonologiques et morphologiques et l'apprentissage de la lecture. Il est intéressant de constater que 46% de la variance de la lecture est prédite par ses deux variables. Les connaissances morphologiques ont une infuence directe significative sur le niveau de lecture en 2ème année. Toutefois, les connaissances phonologiques n'affectent pas directement les performances en lecture puisque la relation linéaire entre les deux dimensions est non significative. Le modèle élaboré a permis de mettre en évidence l'effet médiateur des connaissances phonémiques sur la relation entre les connaissances phonologiques à T1 et l'apprentissage de la lecture à T3. Cet effet de médiation est total puisque l'effet indirect est significatif tandis que l'effet direct entre les deux variables est non significatif.

Ce modèle confirme l'importance des connaissances morphologiques dans l'apprentissage de la lecture en arabe. Si les connaissances phonologiques sont primordiales pour cet apprentissage, les connaissances morphologiques pourraient l'être tout autant. La dimension morphologique serait donc un élément à prendre en compte, notamment dans le repérage précoce des enfants « à risque ».

L'adéquation du modèle proposé aux données recueillies peut être jugée acceptable. La qualité de l'ajustement entre les données et le modèle théorique est confirmée par les coefficients de détermination significatifs des variables dépendantes et par l'existence d'un ensemble de relations de causalité significatives directes entre les variables.

Evaluation des enseignants et prédiction de la réussite dans l'apprentissage de la lecture

Dans cette recherche, nous avons essayé de montrer le rôle et l'apport des évaluations réalisées par les enseignants dans la prédiction de la réussite ou des difficultés d'apprentissage de la lecture. Les résultats montrent que les enseignants de CP ont un rôle important à jouer dans le repérage des difficultés des élèves. En effet, l'évaluation des comportements et compétences scolaires effectuée par les enseignants de CP pourrait être utilisée de manière complémentaire aux évaluations langagières et cognitives dans une perspective de prévention des difficultés d'apprentissage (Quay & Steele, 1998 ; Guimard & Florin, 2008). De plus, ces évaluations sont significativement corrélées avec les performances académiques ultérieures. L'étude longitudinale de Guimard, Gosnefroy & Florin (2007) montre qu'un questionnaire d'évaluation renseigné par les enseignants en début de CP explique significativement les performances en français et en mathématique et le parcours scolaire des élèves en CE2 et en 6ème année. Ce pouvoir explicatif reste significatif et indépendant des prédictions réalisées à partir des évaluations standardisées de CP et des caractéristiques socio-démographiques des élèves.

Les résultats de notre étude montrent, qu'en contexte tunisien, de telles évaluations ont une bonne valeur prédictive du niveau de lecture ultérieur. Certains comportements des élèves en classe évalués par les enseignants semblent constituer des indicateurs pertinents pour identifier dès la CP les enfants en difficulté en lecture. En effet, plusieurs items du questionnaire sont significativement liés aux performances en lecture. Nous constatons que les items comportementaux, plus particulièrement l'organisation en classe, le suivi du rythme et la confiance en soi, sont assez fortement corrélés aux scores ultérieurs en identification de mots. Le niveau de compétence scolaire estimé par les enseignants semble également prédire leurs performances en lecture. Toutefois, les meilleures corrélations sont obtenues avec le score global au questionnaire quelle que soit l'épreuve de lecture considérée et quel que soit le niveau scolaire.

Les analyses de régression entre les items du questionnaire et l'épreuve d'identification de mots en 1ère année indiquent que l'organisation en classe et le respect des règles apportent une contribution significative non négligeable à l'explication de la variance en lecture. Plusieurs items du questionnaire apportent également une contribution significative à l'explication des performances en lecture en 2ème année quelle que soit la variable dépendante considérée. Nous notons en particulier la contribution de l'organisation en classe, du niveau du vocabulaire, de la confiance en soi et de la quantité de participation en classe.

Deux analyses de régression pas à pas ont été conduites afin de comparer le pouvoir prédictif du score global au questionnaire et celui des compétences langagières mesurées en CP (variables indépendantes) sur les performances en lecture en 1ère année et en 2ème année. Les résultats ont montré que le score global au questionnaire n'apporte pas de contribution significative à la variance des performances en identification de mots en 1ère année. En revanche, les évaluations des enseignants constituent de meilleurs prédicteurs du niveau de lecture en 2ème année (IME2 et Timé2) que les mesures langagières. Ce type d'outil pourrait donc être utilisé dans le cadre d'actions de prévention des difficultés scolaires de manière complémentaire aux évaluations langagières et cognitives.

Le questionnaire d'évaluation des compétences et comportements scolaires renseigné par les enseignants de CP constitue un prédicteur puissant des progrès futurs en lecture quelle que soit l'épreuve de lecture utilisée. Ces résultats sont en concordance avec les données de la littérature qui ont montré que les comportements des élèves en classe décrits par les enseignants et leurs évaluations des performances scolaires des élèves constituent des indicateurs particulièrement pertinents pour différencier les élèves et prédire leur évolution (Florin, 1991 ; Florin, Guimard & Nocus, 2002 ; Guimard & Florin, 2007 ; Guimard, Cosnefroy & Florin, 2007 ; Quay & steele, 1998). Selon les difficultés renseignées par les enseignants dans ce type de questionnaire et avec des évaluations standardisées complémentaires si nécessaire, des aides spécifiques peuvent être mises en place précocement. Comme le suggèrent Guimard & Florin (2007), les évaluations effectuées par les enseignants et les évaluations cognitives et langagières devraient être utilisées de manière complémentaire. Les premiers constituant une procédure économique proposée dans le premier temps du repérage et les seconds sont nécessaires pour confirmer ou infirmer cette première évaluation et préciser, le cas échéant, l'ampleur et la nature des difficultés des élèves.

En Tunisie, l'introduction de la classe préparatoire dans le système scolaire est assez récente. Il est donc nécessaire de développer la culture de l'évaluation des enseignants en Tunisie pour une meilleure prévention des difficultés scolaires des enfants. Un des problèmes rencontrés par les enseignants est celui de la gestion de l'hétérogénéité des élèves. Certains enfants savent lire, d'autres ne savent pas encore identifier les lettres et écrire leur prénom. De plus, les difficultés des élèves sont le plus souvent perçues en termes trop généraux : problèmes familiaux, psychologiques, manque de maturité, hyperactivité, etc. Les enseignants se trouvent dans l'incapacité de répondre seuls à ces problèmes et ne disposent pas des moyens nécessaires pour aider les enfants en difficulté. Les recherches doivent ainsi contribuer à la formation des enseignants à la prise en charge des difficultés des élèves compte tenu de l'importance des enjeux pour leur réussite scolaire.

Ces résultats permettent d'établir un cadre théorique qui a des implications sur l'intervention scolaire et clinique. Ils permettront d'orienter la recherche vers la prévention en considérant l'évaluation des enseignants comme une source d'information à prendre en compte dans un premier niveau de repérage des difficultés des élèves. Des travaux complémentaires contrôlant différentes variables motivationnelles et contextuelles qui sont susceptibles d'influencer de façon significative sur leurs évaluations des comportements des élèves semblent particulièrement requis. Les évaluations de l'enseignant relatives à certaines dimensions conatives des élèves semblent indispensables à analyser (e.g., attentes d'effort, d'autonomie, estime de soi, etc.).

Valeur prédictive individuelle des différentes mesures

Un dernier objectif de cette recherche était d'étudier la valeur prédictive individuelle des évaluations langagières et cognitives au niveau de la CP et de la 1ère année et du questionnaire renseigné par les enseignants dans une perspective de prévention des difficultés d'apprentissage. De nombreux travaux soulignent l'intérêt d'une prévention précoce des difficultés scolaires des élèves (Calgar, 1989 ; Guimard & Florin, 2008 ; Zorman & al., 2004). Une intervention préventive peut avoir lieu, dès l'école maternelle, auprès d'enfants dont diverses compétences langagières précocement déficitaires laisseraient entrevoir des difficultés ultérieures dans l'apprentissage de la langue écrite. Un second type d'intervention peut être mis en place dès les premières années de l'école primaire avec une aide massive et ponctuelle auprès d'apprentis lecteurs en difficultés avérées. C'est donc dès les premières années que l'on doit d'abord prévenir l'échec dans l'apprentissage de la lecture et la

démobilisation scolaire qui en résulte souvent. De façon générale, les travaux qui se sont préoccupés de mesurer les prédictions individuelles de réussite ou de difficulté montrent que les évaluations standardisées et les évaluations réalisées par les enseignants prédisent mieux la réussite que les difficultés (Elbro & Peterson, 1999 ; Guimard & Florin, 2007 ; Piquard-kipffer, 2003 ; Zorman & al., 2004).

Dans le but d'étudier la valeur prédictive individuelle des différentes mesures, nous avons constitué trois scores composites : les capacités langagières à T1, les capacités langagières à T2 et les capacités cognitives à T1. Les tests statistiques utilisés ont permis de calculer la sensibilité, la spécificité, la valeur prédictive positive (VPP) et la valeur prédictive négative (VPN) de ces différentes capacités. Les résultats montrent que les mesures de CP et de 1ère année parviennent à mieux prédire la réussite ultérieure en lecture que les difficultés. Les évaluations de CP ont permis de prédire plus de 78% de la réussite des élèves en lecture une année plus tard. Toutefois, ces évaluations ne parviennent à prédire les risques d'échec en lecture que dans un cas sur deux (sensibilité et VPP autour de 50%). Les meilleures prédictions individuelles du niveau de lecture en 1ère année sont obtenues par le questionnaire renseigné par les enseignants qui parvient à repérer 55% des enfants en difficulté. Les plus faibles prédictions parviennent du score cognitif (sensibilité de 43% et VPP de 39%).

La valeur prédictive individuelle des différents indicateurs de CP reste stable deux années plus tard. Ainsi, chacune des dimensions de CP et de 1ère année prédit bien la réussite puisque les spécificités et les VPP sont comprises entre 78% et 91% et ceci quelle que soit l'épreuve de lecture considérée. Cependant, moins de 60% des enfants en difficultés sont identifiés par ces différentes évaluations. Le questionnaire aux enseignants garde sa première place puisque plus de 60% des élèves supposés en difficulté par les enseignants de CP le sont effectivement deux ans plus tard en lecture. Les prédictions individuelles à partir des scores langagiers à T1 et T2 sont très proches (sensibilités et VPP se situant autour de 50%). Nous constatons également que la prédiction individuelle à partir des épreuves cognitives est meilleure lorsque l'indicateur de difficulté ou de réussite en lecture est déterminé à partir des épreuves de lecture en 2ème année.

Les résultats de l'analyse prédictive individuelle montre combien il est délicat de choisir un moment pour évaluer les performances des enfants et d'attribuer très tôt une étiquette de « bon » ou de « faible » lecteur. En effet, nous avons relevé la présence de « faux positifs » c'est-à-dire d'enfants qui en CP avaient un niveau moyen ou bon d'habiletés langagières et

cognitives et dont les performances en lecture étaient faibles deux ans plus tard et de « faux négatifs » qui au contraire faisaient preuve de bonnes capacités en lecture alors qu'ils étaient en difficulté deux ans plus tôt. Les mesures effactuées en CP et en 1ère année ne peuvent avoir une forte sensibilité, bien qu'elles rendent compte de plus de la moitié des risques d'échec de l'apprentissage de la lecture. Certes, tous les enfants qui présentent dès la CP des difficultés dans certains domaines, ne deviendront pas tous des faibles lecteurs. Les premiers contacts avec l'écrit leur permettront de rattraper une partie de leur retard. Comme le souligne Zazzo (1978), une évaluation prédictive doit intégrer la dynamique du développement et l'importance des transformations cognitives et comportementales chez les enfants de ces niveaux d'âge. A cela s'ajoutent les différences de fonctionnement et d'exigences entre la CP et l'école primaire qui peuvent expliquer par exemple que la prédiction soit meilleure lorsqu'elle concerne les futurs bons lecteurs que les futurs faibles lecteurs (Guimard & Florin, 2007).

Les analyses différentielles rétrospectives ont permis d'évaluer l'effet à court terme des différences individuelles observées en CP. Les enfants ont été regroupés en trois niveaux (bons, moyens et faibles) en fonction de leurs résultats aux épreuves de lectures en 2ème année. Les comparaisons effectuées en CP montrent des différences importantes entre les élèves forts et faibles en lecture dans les mesures les plus liées à l'apprentissage de la lecture. De même, les faibles lecteurs ont été évalués moins positivement par les enseignants que les bons lecteurs. En CP, les futurs faibles lecteurs sont nettement plus nombreux à avoir de faibles performances dans différentes composantes du langage oral, à l'exception de la détection de l'intrus de rime et de la dénomination rapide. Le résultat qui nous intéresse le plus est le fait que les faibles lecteurs de CP tendent à être toujours les plus faibles lecteurs deux ans plus tard. Ceci suggère qu'un ciblage des difficultés des élèves est plus aisé au début de la scolarité primaire qu'à son terme et que les interventions pédagogiques spécifiques doivent être précoces pour éviter que les difficultés d'apprentissage ne s'installent et concourent à placer l'élève en situation d'échec. Les conséquences de ces difficultés sont lourdes, le faible lecteur devient vite mauvais élève. La difficulté d'un enfant doit être décelée dès qu'un signe de défaillance apparait afin que les actions appropriées soit mises en œuvre aussitôt. Le succès de la remédiation dépend ainsi de la précocité des interventions.

Au terme de ces développements, nous pouvons maintenant faire le point sur les contributions de cette étude. Elles se situent à plusieurs niveaux :

Sur le plan théorique, cette recherche, est à notre connaissance, la première à avoir entamé une étude prédictive de l'apprentissage de la lecture en langue arabe. Les problèmes soulevés dans cette thèse sont nés d'un souci en tunisie de prévenir l'échec scolaire et l'abondan scolaire qui en découle et la création d'outils adaptés susceptibles de repérer les enfants « à risque ». Le dépistage précoce des difficultés d'apprentissage de la lecture nécessite de savoir quelles sont les les processus ou les pré-requis impliqués dans cet apprentissage. Les résultats obtenus confirment la pertinence théorique d'aller au-delà des données obtenues dans d'autres langues et d'explorer plus avant la contribution de certaines contraintes linguistiques (caractéristiques de la langue arabe et du contexte d'apprentissage) afin de clarifier le rôle des facteurs socioculturels dans le développement des compétences en lecture. Cette recherche a permis de mieux apprécier ce qui relève de processus généraux et ce qui dépend des spécificités des langues ou des systèmes d'écriture. Les caractéristiques de la langue d'apprentissage est un facteur important qu'il est nécessaire de prendre en compte dans les études prédictives.

Par ailleurs, la présente étude ouvre des possibilités d'applications au plan éducatif. Les outils relativement simples utilisés ici pour évaluer les compétences langagières et cognitives des enfants pourraient facilement être intégrés à des batteries d'évaluation pour l'identification précoce et la prévention des difficultés d'apprentissage de la langue écrite. Le questionnaire renseigné par les enseignants ajoute une information pertinente, complémentaire aux évaluations standardisées. En utilisant une adaptation du test d'identification de mots Timé2 (Ecalle, 2003), nous avons souhaité dégager des éléments de réflexion susceptibles de permettre la construction ultérieure d'un instrument plus élaboré et plus fin qui sera probablement étalonné à la population tunisienne.

Un autre aspect original de ce travail nous semble mériter d'être mentionnée. Il s'agit de la méthode de traitement des données proposée : analyses en composantes principales avec rotations, corrélations entre facteurs, analyses de régressions multiples, valeur prédictive individuelle. En adoptant une étude longitudinale nous avons pu conférer à ce travail une méthodologie rigoureuse. De plus, le recours aux techniques modernes des modèles d'équations structurelles constitue un élément de qualité contribuant à un traitement original et en profondeur des données. Les analyses factorielles confirmatoires, qui constituent une originalité de l'étude car elles sont encore peu utilisées dans ce domaine, a permis de vérifier la validité de construits psychologiques en établissant des relations entre les variables latentes (compétences supposées) et leurs indicateurs.

La présente étude, comme tout travail de recherche, comporte toutefois un certain nombre de limites et soulève de nouvelles questions. Deux limites en particulier méritent une attention particulière. Premièrement, le rôle des facteurs sociodémographiques (âge, sexe, niveau socio-économique) et conatifs a été négligé dans cette étude. A l'entrée à l'école, les écarts de performances sont toujours très marqués en fonction de l'âge et de l'origine sociale. Les élèves nés en début d'année civile obtiennent de meilleures performances scolaires que ceux nés en fin d'année et ont une probabilité plus élevée de suivre une scolarité sans redoubler (Florin, Cosnefroy & Guimard, 2004).

Dès l'école maternelle, des différences importantes existent entre les élèves sur le plan de leurs performances cognitives. Les habiletés liées au langage oral ont été mises en lien avec d'autres types d'expériences sociales. L'origine sociale joue un rôle significatif puisque, à niveau initial comparable, les enfants originaires des milieux sociaux les plus favorisés progressent davantage. Zorman (1999) a constaté que le niveau de conscience phonologique en grande section de maternelle était fortement déterminé par le niveau de diplôme de la mère et le niveau de langage et de vocabulaire des enfants. Ainsi, plus l'enfant a été élevé dans un milieu culturellement défavorisé, où la pratique langagière est fruste ou peu développée, moins il a l'occasion d'acquérir et d'exercer les connaissances et compétences linguistiques et métalinguistiques nécessaires à l'apprentissage de la lecture. Il existe également des différences de performances entre filles et garçons, à l'avantage des filles. Ces résultats sur les écarts d'acquisitions présents à l'entrée à l'école élémentaire nous conduisent à souligner l'aspect cumulatif des écarts liés aux différentes variables. A cela s'ajoutent les différences de fonctionnement et d'exigences entre l'école élémentaire et l'école primaire. Ainsi, une évaluation prédictive doit intégrer la dynamique du développement et l'importance des transformations comportementales chez les enfants de ces niveaux d'âge.

N'oublions pas par ailleurs que l'apprentissage de la lecture relève aussi de facteurs motivationnels et affectifs. Ainsi, le degré de motivation de l'enfant à apprendre à lire constitue un facteur important dans la réussite de cet apprentissage. La prise en compte des dimensions conatives (satisfaction éprouvée par les enfants, estime de soi, etc.) paraît indispensable pour mieux comprendre le fonctionnement des élèves et analyser leurs difficultés en vue d'un diagnostic et une prise en charge adéquate (Guimard & Florin, 2008). C'est donc dans ce contexte global que l'école maternelle peut agir pour réduire les inégalités cognitives et sociales entre élèves. La faiblesse des connaissances préscolaires qui compromet fortement le déroulement de l'apprentissage alerte sur la nécessité pour l'école de réduire

l'hétérogénéité constatée dans ce domaine de connaissance. L'idée étant d'identifier les compétences scolaires les plus prédictives de la réussite ultérieure pour pouvoir agir sur leur développement très tôt dans la scolarité. Les études futures devront en outre poursuivre l'examen de la contribution des facteurs socioculturels en général et de la diversité des opportunités sociales en particulier et considérer d'autres dimensions des compétences en lecture.

Une deuxième limite est liée au fait que certaines compétences sont très peu développées avant un certain âge. Ainsi, nous n'avons pas pu évaluer le poids des connaissances phonémiques et des connaissances morpho-dérivationnelles en CP puisque ces deux dimensions ne sont pas assez développées à ce niveau scolaire. Avant l'apprentissage de la lecture, les connaissances phonémiques des enfants, comparativement à leurs capacités d'analyse syllabique, sont faibles. De fait, comparer le pouvoir prédictif de ces deux capacités est difficile, les évaluations n'ayant pas été faites au même moment. Le même problème est posé en ce qui concerne les connaissances morphologiques dérivationnelles et flexionnelles. Une question importante concerne la mise en relation des compétences identifiées aux différents niveaux scolaires et leur évolution dans le temps. L'importance relative des différentes compétences citées (conscience phonologique, conscience morphologique, connaissance des lettres, etc.) en tant que prédicteurs des performances ultérieures de lecture peut changer graduellement au cours du développement. Autrement dit, les mesures dans certains domaines sont prédictives de la réussite à un moment donné du développement. Muter & al. (1998) démontrent que les connaissances morpho-syntaxiques jouent un rôle significatif dans l'identification de mots en contexte à l'âge de 9 ans. En revanche, ces connaissances ne sont pas aussi importantes dans la prédiction de la lecture à une étape précoce (avant 5 ans). Dans le même sens, l'étude longitudinale de Casalis & Alexandre (2000) avec des enfants français montre que la conscience phonologique explique la plus grande part de variance explicative de la réussite en lecture à l'âge de 6,5 ans mais lorsque les enfants ont 7,5 ans, la conscience phonologique et les compétences morphologiques apportent chacune une contribution indépendante et significative à l'apprentissage de la lecture. Dans cette perspective, nous souhaitons poursuivre cette étude avec les mêmes enfants pour étudier

la valeur prédictive à long terme des différentes mesures langagières et cognitives et des évaluations des enseignants réalisées en CP[96].

À l'issue de ce travail et conformément aux connaissances établies, le repérage des difficultés dans l'apprentissage de la lecture doit être effectué le plus tôt possible. Les évaluations langagières et cognitives sont des instruments indispensables à une meilleure compréhension du fonctionnement de l'apprenant, au repérage et l'analyse des difficultés que rencontrent les élèves en situation d'apprentissage. Les enseignants peuvent y contribuer et une grande responsabilité leur revient dans ce domaine. Le dépistage précoce permet de déboucher sur des activités de prévention pédagogiques spécifiques, d'adaptation et de remédiation ainsi que les prises en charge rééducatives (Zorman & al., 2004).

Les travaux de la littérature internationale montrent qu'un entraînement en milieu scolaire améliore les compétences des faibles lecteurs (voir INSERM, 2007 pour une synthèse). L'intérêt des entraînements tient à la nécessité d'éviter chez les enfants faibles lecteurs le décalage avec leurs pairs bons lecteurs. Ainsi, les habiletés métalinguistiques seraient susceptibles d'être stimulées par un entraînement spécifique. Une intervention précoce visant à développer la conscience phonologique, en particulier celle des phonèmes, dès la classe préparatoire peut contribuer à réduire le nombre d'enfants en difficulté dans l'apprentissage de la lecture. L'apprentissage explicite du nom des lettres constitue en lui-même un entraînement favorisant le développement des habiletés phonologiques en activant les capacités de traitement métaphonologique du phonème lorsque celui-ci correspond à une lettre connue (Biot-Chevrier & al., 2009). D'autres études récentes montrent également l'intérêt particulier des outils informatisés pour les entraînements phonologiques (Magnan & Ecalle, 2006). En langue arabe, il pourrait être également intéressant de travailler sur la dimension morphologique pour aider à son développement précoce par des entraînements spécifiquement axés sur la morphologie dérivationnelle. Différentes recherches évoquent, en effet, une influence d'un entraînement portant sur la morphologie sur le développement des habiletés phonologiques (Casalis & Colé, 2009). Néanmoins, la spécificité de la langue arabe et de notre système éducatif nécessite des études compléméntaires pour prouver les effets et les limites d'entraînements similaires à ceux de la littérature internationale.

[96] Les enfants de notre échantillon seront scolarisés en 5ème année du primaire au cours de l'année scolaire 2009-2010.

Le dernier centre d'intérêt consiste à envisager un programme d'aide pour les enfants en début de l'apprentissage de la lecture dans un objectif de prévention. Il est primordial de détecter le plus rapidement possible les futurs faibles lecteurs afin de construire des projets d'aide adaptés. En effet, certains enfants abordent les premières années de scolarisation en situation déjà problématique. C'est donc dès la classe préparatoire que l'on doit prévenir l'échec dans l'apprentissage de la lecture afin d'améliorer le rendement et la rentabilité du système éducatif tunisien.

En tunisie, nous participons actuellement à un grand projet de recherche ayant pour objectif de contribuer au dépistage précoce des troubles spécifiques des apprentissages (troubles du langage oral et écrit et troubles des apprentissages numériques) et de participer à l'élaboration d'un diagnostic en vue de définir une stratégie de prise en charge. Ce projet s'inscrit dans une démarche de prévention visant à repérer dès la CP les enfants présentant des difficultés, à leur apporter, au sein de l'école, une aide établie élaborée en partenariat par les enseignants et les orthophonistes, et à les orienter éventuellement vers des bilans complémentaires et vers des rééducations.

Les perspectives de recherche et de réflexion dans le domaine des difficultés de l'apprentissage de la lecture apparaissent plus que jamais auparavant intégrantes, larges et prometteuses.

Bibliographie

Abu-Rabia, S. (1995). Learning to read in Arabic: Reading, syntactic, orthographic and working memory skills in normally achieving and poor Arabic readers. *Reading Psychology: An International Quarterly, 16,* 351-394.

Abu-Rabia, S. (1996). The role of vowels and context in the reading of highly skilled native arabic readers. *Journal of Psycholinguistic Research, 6,* 629-641.

Abu-Rabia, S. (1997). Reading in Arabic orthography: the effect of vowels and context on reading accuracy of poor and skilled native arabic readers in reading paragraphs, sentences, and isolated words. *Journal of Psycholinguistic Research, 4,* 465-482.

Abu-Rabia, S. (1998). Reading Arabic texts: Effects of text type, reader type and vowelisation. *Reading and Writing: An Interdisciplinary Journal, 10,* 105-119.

Abu-Rabia, S. (1999). The effect of Arabic vowels on the reading comprehension of second- and sixth-grade native Arab children. *Journal of Psycholinguistic Research, 1,* 93-101.

Abu-Rabia, S. (2000). Effects of exposure to literary arabic on reading comprehension in a diglossic situation. *Reading and Writing, 13*(1), 147-157.

Abu-Rabia, S. (2001). The role of vowels in reading semitic scripts : Data from Arabic and Hebrew. *Reading and Writing: An Interdisciplinary Journal, 14,* 39-59.

Abu-Rabia, S. (2007). The Role of Morphology and Short Vowelization in Reading Arabic among Normal and Dyslexic Readers in Grades 3, 6, 9, and 12. *Journal of Psycholinguistic Research, 36 (2),* 89-106.

Abu-Rabia, S., & Siegel, L. S. (1995). Different orthographies, different context effects: the effects of Arabic sentence context on skilled and poor readers. *Reading Psychology, 16,* 1-19.

Abu Rabia, S., Share, D., & Mansour, M. S. (2003). Word recognition and basic cognitive process among reading-disabled and normal readers in Arabic. *Reading and Writing: An Interdisciplinary Journal, 16,* 423-442.

Alegria, J., Leybaert, J., & Mousty. Ph. (1994). Acquisition de la lecture et troubles associés. In J. Grégoire & B. Pierart (Eds.), *Evaluer les troubles de lecture : les nouveaux*

modèles théoriques et leurs implications diagnostiques (pp. 105-126). Bruxelles, De Boek Université.

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 de la lecture et de l'écriture chez l'enfant et l'adulte* (pp. 81-96). Marseille : Editions Solal.

Alegria, J., & Morais, J. (1989). Analyse segmentale et acquisition de la lecture. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques* (pp. 173-196). Paris, Neuchâtel, Delachaux & Niestlé.

Alegria, J., Pignot, E., & Morais, J. (1982). Phonetic analysis of speech and memory codes in beginning readers. *Memory and Cognition, 10*, 451-456.

Alegria, J., Marin, J., Carrillo., & Mousty, Ph. (2003). Les premiers pas dans l'acquisition de l'orthographe en fonction du caractère profond ou superficiel du système alphabétique : comparaison entre le français et l'espagnol. In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 51-79). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Ammar, M. (1997). *Les stratégies d'identification de mots écrits en arabe*. Thèse de doctorat en sciences de l'éducation non publiée, Université de Nantes.

Ammar, M. (2002). L'assemblage phonologique : sa nature et son fonctionnement chez des enfants lecteurs en arabe. *Enfance, 2 (54),* 155-168.

Ammar, M. (2003). Le fonctionnement de l'assemblage phonologique chez des enfants lecteurs en arabe. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 96-79). Presses Universitaires de Rennes & Centre de Publication Universitaire De Tunis.

Attia, A. (1966). Différents registres de l'emploi de l'arabe en Tunisie. *Revue Tunisienne des Sciences Sociales, 8*, 115-150.

Baccouche, T. (1994). *L'emprunt en arabe moderne.* Tunis, Académie Tunisienne des Sciences, des Lettres et des Arts et l'Institut Bourguiba des Langues Vivantes.

Baccouche, T. (2003). La langue arabe : spécificités et évolution. In M. Romdhane, J.-E. Gombert & M. Belajouza (Eds), *L'apprentissage de la lecture : perspectives*

comparatives (pp. 377-387). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Backman, J., Bruck, M., Hebert, M., & Seidenberg, M. S. (1984). Acquisition and use of spelling sound correspondences in reading. *Journal of Experimental Child Psychology, 38*, 114-133.

Baddeley, A. D. (1986). *Working memory*. Oxford : Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *recent advances in learning and motivation* (Vol. 8, pp. 47-89). Academic Press, New York.

Ball, E. W., & Blachman, B. A. (1991). Does phoneme segmentation training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26,* 49-66.

Baron, J., & Treinman, R. (1980). Some problems in the study of differences in cognitive processes. *Memory and Cognition, 4,* 313-321.

Bastien-Toniazzo, M. (1992). La représentation du mot écrit aux débuts de la lecture. *L'Année Psychologique, 92,* 489-509.

Bastien-Toniazzo, M., & Jullien, S. (2001). Nature and importance of the logographic phase in learning to read. *Reading and Writing, 14 (1-2),* 119-143.

Bastien-Toniazzo, M., & Bastien, C. (2003). Considérations théoriques sur l'acquisition de la lecture et implications pédagogiques. In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds), *L'apprentissage de la lecture: perspectives comparatives* (pp. 321-335). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Beauvillain, C. (1990). Le rôle de la structure morphologique dans la reconnaissance visuelle des mots, in R. Kolinsky, J. Morais & J. Segui. *La reconnaissance des mots dans les différentes modalités sensorielles*. Paris, PUF.

Belajouza, M. (2003). Stratégies d'identification des mots, compétences phonologiques et morphologiques des lecteurs en difficulté. In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 195-211). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Belajouza, M., & Snoussi, F. (2006). Apprentissage de la lecture en langue arabe et contraintes externes. *Muqabasat Intersignes, 1,* 91-113.

Belkhadi-Maaouia, A. (1988). Le passage du dialectal à l'arabe littéral chez les enfants tunisiens à leur entrée à l'école. *Cahier du C.E.R.E.S, série psychologique, 5*, 119-192.

Belkhadi-Maaouia, A. (2003). La compréhension de l'énoncé et la structure syntaxique de la phrase arabe. In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 143-160). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Ben dror, I. Bentin, S., & Frost, R. (1995). Semantic, phonological and morphologic skills in reading disabled and normal children: evidence from perception and production of spoken Hebrew. *Reading Research Quarterly, 34*, 876-893.

Ben Fadhel, S. E. (2003). Apprendre à lire et écrire: une construction socio-cognitive. In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 281-295). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Ben Rejeb, R. (2001). *Intelligence, test et culture*. *Le contexte tunisien*. L'Harmattan.

Bentin, S., & Ibrahim, R. (1996). New evidence for phonological processing during visual word recognition: the case of Arabic. *Journal of Experimental Psychology: Learning, Memory and Cognition, 2,* 309-323.

Bentin, S., Hammer, R., & Cahan, S. (1991). The effects of aging and first-grade schooling on the development of phonological awareness. *Psychological Science, 2*, 271-274.

Besse, A. S., Demont, E., & Gombert, J. E. (2007). Effet des connaissances linguistiques en langue maternelle (arabe vs portugais) sur les performances phonologiques et morphologique en français langue seconde. *Psychologie française, 52*, 89–105.

Besse, A. S. (2007). *Caractéristiques des langues et apprentissage de la lecture en langue première et en français langue seconde. Perspective évolutive et comparative entre l'arabe et le portugais*. Thèse de doctorat en psychologie non publiée, Université Rennes 2.

Billard, C., Fluss, J., Ducot, C., Warszawski, J., Ecalle, J., Magnan, A., Richard, G., & Ziegler, J. (2008). Etude des facteurs liés aux difficultés d'apprentissage de la lecture à partir d'un échantillon de 1062 enfants de seconde année d'école élémentaire. *Archive de Pédiatrie, 15(6)*, 1058-1067.

Biot-Chevrier, C., Ecalle, J., & Magnan, A. (2008). Pourquoi la connaissance du nom des lettres est-elle si importante dans l'apprentissage de la langue écrite ? *Revue Française de Pédagogie, 162*, 15-27.

Blachère, R., & Gaudefroy-Demombynes, M. (1975). *Grammaire de l'arabe classique*. Paris, Maisonneuve & Larose.

Blanc, N. (2009). *Lecture et habiletés de comprehension chez l'enfant*. Dunod.

Bloodgood, J. W. (1999). What's in a name ? Children's name writing and literacy acquisition. *Reading Research Quarterly, 34(3),* 342-367.

Boukadida, N. (2008). *Connaissances phonologiques et morphologiques dérivationnelles et apprentissage de la lecture en arabe (Etude longitudinale).* Thèse de Doctorat en psychologie non publiée, Université de Rennes 2 & Université de Tunis.

Boukadida, N., Gombert, J.E., & Maaouia, A. (2009). Contribution des connaissances morphologiques à l'apprentissage de la lecture en arabe. In N. Marec-Breton, A-S. Besse, F. de la Haye, N. Bonneton & E. Bonjour (Eds.), *L'apprentissage de la langue écrite : Approche cognitive* (pp. 107-118). Rennes : PUR.

Boehm, A. E. (1990). *Test des concepts de base de Boehm : version pré-scolaire*. Paris : ECPA.

Bosse, M. L. (2005). De la relation entre acquisition de l'orthographe lexicale et traitement visuo-attentionnel chez l'enfant. *Rééducation orthophonique, 222*, 9-30.

Bosse, M. L., & Valdois, S. (2009). Influence of the visual attention span on child reading performance: A cross-sectional study. *Journal of Research in Reading, 32 (2),* 230-253.

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia : the visual attention span deficit hypothesis. *Cognition, 104*, 2,198-230.

Bosse, M. L., Valdois, S., & Dompnier, B. (2009). Acquisition du langage écrit et empan visuo-attentionnel : une étude longitudinale. In N. Marec-Breton, A-S. Besse, F. de la Haye, N. Bonneton & E. Bonjour (Eds.), *L'apprentissage de la langue écrite : Approche cognitive* (pp. 167-178). Rennes : PUR.

Boulc'h, L. (2008). *Implication du contrôle excécutif dans le décodage en lecture. Etude chez l'enfant et l'adolescent*. Thèse de Doctorat en psychologie non publiée, Université d'Angers.

Boulc'h, L., Gaux, C., & Boujon. (2007). Implication des fonctions exécutives dans le décodage en lecture : étude comparative entre normolecteurs et faibles lecteurs de CE2. *Psychologie Française, 52(1),* 71-87.

Bryant, P., Bradley, L., MacLean, M., & Crossland, J. (1989). Nursery rhymes, phonological skills and reading. *Journal of Child Language, 16,* 407-428.

Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpho-lexical reading develop in readers of a shallow orthography ? *Brain and Language, 81,* 568-586.

Byrne, B., & Barnesley, F. R. (1990). Acquiring the alphabetic principle: a case for teaching recognition of phoneme identity. *Journal of Educational Psychology, 82,* 805-812.

Byrne, B., Freebody, P., & Gates, A. (1992). Longitudinal data on the relation of word-reading strategies to comprehension, reading time and phonemic awareness. *Reading Research Quarterly, 27,* 141-151.

Calgar, H. (1989). *La psychologie scolaire*. Paris : PUF.

Calvet, L. J. (1993). *La sociolinguistique*. Paris, PUF.

Cantineau, J. (1960). *Etude de linguistique arabe*. Paris, Klincksieck.

Carayon, C. (1991). Lire : pourquoi faire ? La représentation de la lecture chez l'enfant. *Bulletin de Psychologie, 400,* 215-222.

Carlisle, J. F. (1995). Morphological awaness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189-209). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Carlisle, J. F., & Nomanbhoy, D. M. (1993). Phonological and morphological awareness in first graders. *Applied psycholinguistics, 19 (2),* 177-195.

Caron, J. (1989). *Précis de psycholinguistique*. Paris, PUF.

Casalis, S. (1995). *Lecture et dyslexie de l'enfant*. Presse Universitaire du Septentrion.

Casalis, S., & Colé, P. (2009). On the relationship between morphological and phonological awareness: Effects of training in kindergarten and in first-grade reading. *First Language, 29(1),* 113-142.

Casalis, S., & Louis-Alexandre, M. F. (2000). Morphological analysis, phonological analysis and learning to read Frensh: a longitudinal study. *Reading and writing, 12,* 303-335.

Casalis, S., & Lecoq. (1992). La dyslexie développementale. In M. Fayol, J. E. Gombert, P. Lecoq, L. Sprenger-Charolles & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 195-237). Paris, PUF.

Casalis, S., Colé, P., & Royer, C. (2003). Traitement morphologique et dyslexie: une stratégie compensatoire pour les dyslexiques? *Glossa, 85*, 4-17.

Castel, C., Pech-Georgel, C., George, F., & Ziegler, J. C. (2008). Lien entre dénomination rapide et lecture chez les enfants dyslexiques. *L'Année Psychologique, 108*, 395-422.

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 : PUF.

Colé, P., & Valdois, S. (2007). L'apprentissage de la lecture et ses troubles. In A. Blaye & P. Lemaire (Eds.), *Le développement cognitif de l'enfant* (pp. 187-222). Editions De Boeck.

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., Royer, C., Leuwers, Ch., & Casalis, S. (2004). Les connaissanes morphologiques dérivationnelles et l'apprentissage de la lecture chez l'apprenti lecteur français du CP au CE2. *L'Année Psychologique, 107*, 701-750.

Colé, P., Magnan, A., & Grainger, J. (1999). Syllable-sized units in visual word recognition: evidence from skilled and beginning readers of French. *Applied Psycholinguistics, 20*, 507-532.

Content, A. (1993). *Le rôle de la médiation phonologique dans l'acquisition de la lecture, lecture-écriture : acquisition*. Paris, Nathan.

Content, A. (1996). Modèles de l'acquisition de la lecture : perspectives récentes. In S. Carbonnel & P. Gilly (Eds.), *Approche cognitive des troubles de lecture et de l'écriture chez l'enfant et l'adulte* (pp. 63-81). Solal Editeurs.

Content, A., & Leybeart, J. (1992). L'acquisition de la lecture : influence des méthodes d'apprentissage. In P. Lecoq (Ed.), *La lecture : processus, apprentissage, troubles* (pp. 181-211). Presses Universitaires de Lilles.

Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italien children. *Applied Psycholinguistics, 9,* 1-16.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Defior-Citoler, S., & Herrera-Torres, L. (2003). Les habiletés de traitement phonologique des enfants prélecteurs espagnols. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 161-176). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

De La Haye, F., & Bonneton-Botté, N. (2009). Incidence d'un entraînement à la production d'inférences sur le niveau de compréhension en lecture. In N. Marec-Breton, A-S. Besse, F. de la Haye, N. Bonneton & E. Bonjour (Eds.), *L'apprentissage de la langue écrite : Approche cognitive* (pp. 187-200). Rennes : PUR.

Demont, E. (2003). Conscience phonologique et mémoire de travail : quelle contribution respective aux difficultés en lecture ? In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 212-232). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Demont, E. (2002). *Lecture et processus cognitifs : une approche psycholinguistique de l'apprentissage de la lecture et de ses difficultés*. Thèse d'Habilitation à diriger des recherches, Université Louis Pasteur-Strasbourg.

Demont, E. (1994). Conscience phonologique, conscience syntaxique. Quel(s) rôle(s) dans l'apprentissage efficace de la lecture ? In J. Grégoire & B. Piérart (Eds.), *Evaluer les troubles de la lecture : les nouveaux modèles théoriques et leurs implications diagnostiques* (pp. 195-208). Bruxelles, De Boek Université.

Demont, E., & Gombert, J.-E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology, 66,* 315-332.

Demont, E., & Gombert, J.-E. (2004). L'apprentissage de la lecture : évolution des procédures et apprentissage implicite. *Enfance, 3,* 245-257.

Demont, E., & Gombert, J.-E. (2007). Relations entre conscience phonologique et apprentissage de la lecture: peut-on sortir de la relation circulaire? In E. Demont & M.

N. Metz-Lutz (Eds.), *L'acquisition du langage et ses troubles* (pp. 47-79). Marseille: Solal.

Derguini, M., & Mekideche, T. (2003). Quel(s) modèle(s) d'accès au lexique chez l'apprenti lecteur arabophone disposant de deux codes ? In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 131-160). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory and Cognition, 8,* 195-209.

Dubois, J. (1989). *Dictionnaire de linguistique et des sciences du langage.* Larousse.

Ecalle, J. (2000a). Prédiction de la réussite scolaire en lecture-écriture au cycle II. *Revue Européenne de Psychologie Appliquée, 50 (1),* 81-86.

Ecalle, J. (2000b). Lire et écrire des mots au cours préparatoire : aspects prédictifs. *Revue de Psychologie de l'Education, 1,* 42-58.

Ecalle, J. (2003). *Timé2 : Test d'Identification de mots écrits pour enfants de 6 à 8 ans*. Paris : ECPA.

Ecalle, J. (2004). Les connaissances des lettres et l'écriture du prénom chez l'enfant français avant l'enseignement formel de la lecture-écriture. *Psychologie Canadienne. 45(1),* 111-118.

Ecalle, J., Armand, V., Chevillot-Junker, N., & Orsat, Ch. (2001). Approche cognitive des difficultés d'apprentissage de la langue écrite en contextes éducatifs. *Psychologie & Education, 44,* 71-90.

Ecalle, J., & Magnan, A. (2002). *L'apprentissage de la lecture : fonctionnement et développement cognitifs*. Armand Colin.

Ecalle, J., & Magnan, A. (2003). Traitements épiphonologique et métaphonologique et apprentissage de la lecture. In M. N. Romdhane, J.-E. Gombert & M. Belajouza, *L'apprentissage de la lecture. Perspectives comparatives* (pp. 251-265). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Ecalle, J., & Magnan, A. (2007a). Sensibilité phonologique et apprentissage de la lecture. *Rééducation Orthophonique, 229,* 61-74.

Ecalle, J., & Magnan, A. (2007b). Development of phonological skills and learning to read in French. *European Journal of Psychology of Education, 2*, 153-167.

Ecalle, J., Magnan, A., & Biot-Chevrier, C. (2008). Alphabet knowledge and early literacy skills in French beginning readers. *European Journal of Developmental Psychology, 5(3)*, 303-325.

Ecalle, J., Magnan, A., & Biot-Chevrier, C. (2009). Analyse de l'évolution des connaissances alphabétiques précoces et de leur poids en identification de mots écrits : une étude longitudinale de la GS au CP. In N. Marec-Breton, A.-S. Besse, F. De La Haye, N. Bonneton-Botte, & E. Bonjour (Eds.), *Apprentissage de la langue écrite. Approche cognitive*, (pp. 13-21). Rennes : PUR.

Ecalle, J., Magnan, A., & Bouchafa, H. (2002). Le développement des habiletés phonologiques avant et au cours de l'apprentissage de la lecture : de l'évaluation à la remédiation. *Glossa, 82*, 4-12.

Ecalle, J., Magnan, A., & Ramus, F. (2007). L'apprentissage de la lecture et ses troubles. In S. Ionescu & A. Blanchet (Dir.). *Nouveau cours de psychologie. Psychologie du développement et de l'éducation* (vol. coordonné par J. Lautrey) (pp. 229-275). Paris : PUF.

Ehri, L.C. (1989). Apprendre à lire et écrire des mots. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques* (pp. 103-128). Paris, Neuchâtel, Delachux et Niestlé.

Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read word by analogy. *Reading Research Quarterly, 27*, 12-26.

Elbro, C., & Arnbak, E. (1996). The role of morpheme recognition and morphological awarness in dyslexia. *Annals of dyslexia, 46*, 209-240.

Ellis, N., & Large, B. (1988). The early stages of reading: A longitudinal study. *Applied Cognitive Psychology, 2*, 47-76.

Faure, J., Ecalle, J., & Magnan, A. (2000). *Les procédures analogiques dans la reconnaissance de mots écrits chez des apprentis lecteurs*. Colloque de l'Association de Psychologie Scientifique de Langue Française (APSLF). Nantes, 14-16 septembre.

Fayol, M. (1992). Comprendre ce qu'on lit: de l'automatisme au contrôle. In M. Fayol, J. E. Gombert, P. Lecoq, L. Sprenger-Charolle & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 73-105). Paris, PUF.

Fayol, M., & Gombert, J.-E. (1999). L'apprentissage de la lecture et de l'écriture. In J.A. Rondal et E. Espéret (Eds.), *Manuel de Psychologie de l'enfant* (pp. 565-594). Bruxelles: Mardaga.

Fayol, M., & Jaffré, J. P. (2008). *Orthographier*. Paris : PUF.

Feldman, L. B. Frost, R., & Pnini, T. (1995). Decomposing word into their constituent morphemes : Evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *4*, 947-960.

Fijalkow, J. (1986). *Mauvais lecteurs pourquoi* ? Paris, PUF.

Fitouri, Ch. (1983). *Biculturalisme, bilinguisme et éducation*. Paris, Neuchâtel, Delachaux et Niestlé.

Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. L. (1993). Effects to listening to story reading on aspect of literacy acquisition in a diglossic situation. *Reading Research Quarterly*, *28(1)*, 71-79.

Ferrand, L. (1995). Evaluation du rôle de l'information phonologique dans l'identification des mots écrits. *L'Année Psychologique, 95*, 293-315.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of psychonomic society*, *31*, 119-122.

Ferrand, L., & Grainger, J. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Langage, 33*, 218-233.

Florin, A. (1991). *Pratique du langage à l'école maternelle et prédiction de la réussite scolaire.* Paris, PUF.

Florin, A., Cosnefroy, O., & Guimard, P. (2004). Trimestre de naissance et parcours scolaires. *Revue Européenne de Psychologie Appliquée*, *54*, 237-246.

Florin, A., Guimard, Ph., Khomsi, A., Ecalle, J., & Guégan, J. F. (1999). Maîtrise de l'oral en grande section de maternelle et conceptualisation de la langue écrite en début de cours préparatoire. *Revue Française de Pédagogie*, *126*, 71-82.

Florin, A., Guimard, Ph., & Nocus. I. (2002). Les évaluations des enseignants et la prédiction des compétences langagières de leurs élèves : étude longitudinale à l'école maternelle et à l'école élémentaire. *Le langage et l'homme, 2*, 175-190.

Florin, A., Guimard, Ph., & Nocus. I. (2008). *Présentation, codage et utilisation du questionnaire pour les élèves en maternelle (Q-EM)*. Université de Nantes (Labécd, EA 3259).

Fluss, J., Ziegler, J., Ecalle, J., Magnan, A., Warszawski, J., Ducot, C., Richard, G., & Billard, C. (2008). Prévalence des troubles d'apprentissages du langage écrit en début de scolarité : l'impact du milieu socioéconomique au travers de trois zones d'éducations distinctes. *Archives de Pédiatrie, 15(6)*, 1049-1057.

Foulin, J. N. (2007). La connaissance des lettres chez les pré-lecteurs : aspects pronostiques, fonctionnels et diagnostiques. *Psychologie Française, 4*, 431-444.

Foulin, J. N., & Pacton, S. (2006). La connaissance du nom des lettres précurseur de l'apprentissage du son des lettres. *Education et Francophonie, 2*, 28-55.

Fowler, A. E., & 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, New Jersey: Lawrence Erlbaum Associates.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall & M. Coltheart (Eds.), *surface dyslexia : neuropsychological and cognitive studies of phonological reading* (pp. 301-330). London : Routeledge & Kegan Paul.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological reading in German- and Spanish-speaking children. *Scientific Study of Reading, 2*, 31-54.

Garmadi, S. (1968). La situation linguistique actuelle en Tunisie : problèmes et perspectives. *Revue Tunisienne des Sciences Sociales*, *13*, 13-32.

Georgiou, G. K., Parilla, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages in varying orthographic consistency. *Journal of Educational Psychology, 100*, 566-580.

Gombert, J. E. (1988). La conscience du langage à l'âge préscolaire, *Revue Française de Pédagogie, 83*, 65-81.

Gombert, J. E. (1990). *Le développement métalinguistique*. Paris : PUF.

Gombert, J. E. (1992a). Activités de lecture et activités associées. In M. Fayol, J. E. Gombert, P. Lecoq, L. Sprenger-Charolles & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 107-140). Paris : PUF.

Gombert, J.-E. (1992b). Le rôle des capacités métaphonologiques dans l'apprentissage de la lecture. In A. Bentolila (Ed.), *Les entretiens Nathan*. Editions Nathan.

Gombert, J.-E. (2003). L'apprentissage des codes grapho-phonologiques et grapho-sémantiques en lecture. In M. N. Romdhane, J.E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 19-34). Presses Universitaires de Rennes & Centre de Publication Universitaire De Tunis.

Gombert, J.-E., Bonjour, E., & Marec-Breton, N. (2004). Processus implicites et traitements intentionnels dans l'apprentissage de la lecture. In M. N. Metz, E. Demont, C. Seegmuller, M. de Agostini & N. Bruneau (Eds.), *Développement cognitif et troubles des apprentissages* (pp. 175-192). Solal Editeurs.

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 3 ans* (Vol. 2, pp. 117-150). PUF.

Gombert, J.-E. & Fayol, M. (1995). La lecture-compréhension : fonctionnement et apprentissage. In D. Gaonac'h & C. Golder (Eds.), *Profession enseignant. Manuel de psychologie pour l'enseignement* (pp. 358-381). Paris : Hachette.

Gombert, J.-E., Bryant, P., & Warrick, N. (1997). Les analogies dans l'apprentissage de la lecture et de l'orthographe. In L. Rieben, M. Fayol & C. A. Perfetti (Eds.), *Des orthographes et leur acquisition* (pp. 319-333). Paris : Delachaux & Niestlé.

Gombert, J.-E., Colé, P., Valdois, S., Goigoux, R., Mousty, P., & Fayol, M. (2000). *Enseigner la lecture au cycle 2*. Paris : Nathan Pédagogie.

Gombert, J.-E., Gaux, C., & Demont, E. (1994). Capacités métalinguistiques et lecture : Quels liens ? *Repères, 9*, 47-54.

Goigoux, R. (2000). Apprendre à lire à l'école : les limites d'une approche idéo-visuelle. *Psychologie Française, 45*, 233-243.

Golder, C., & Gaonac'h, D. (1998). *Lire et comprendre : psychologie de la lecture*. Paris: Hachette Education.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child psychology, 42*, 72-83.

Goswami, U. (1988). Orthographic analogies and reading development. *Quarterly Journal of Experimental Psychology, 2*, 239-268.

Goswami, U. (1993). Phonological skills and learning to read. *Ann N Y Acad Sci, 682*, 296-311.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale: Erlbaum.

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. (1998). Children's orthographic representations and linguistic transparency : nonsense word reading in English, Frensh and Spanish. *Applied Psycholinguistics, 19*, 19-52.

Gough, P. B., & Hillinger, M. (1980). Learning to read : An unnatural act. *Bulletin of the Orton Society, 30*, 1-17.

Gough, P. B., & Juel, C. (1989). Les premières étapes de la reconnaissance des mots. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implication pédagogiques* (pp. 88-102). Neuchâtel, Delachaux et Niestlé.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial & Special Education, 7*, 1, 6-10.

Grainger, J., Dichy, J., El-Halfaoui, M., & Bamhamed, M. (2003). Approche expérimentale de la reconnaissance du mot écrit en arabe. In J. P. Jaffré (Eds.), *Dynamique de l'écriture : Approches pluridisciplinaires, revue faits de langue n° 22* (pp. 77-86). Paris: Ophrys.

Guimard, Ph., & Florin, A. (2001). Comportement scolaire en moyenne section de maternelle et prédiction de la réussite scolaire à l'école élémentaire. *Psychologie et Psychométrie, 1*, 75-100.

Guimard, Ph., & Florin, A. (2007). Les évaluations des enseignants en grande section de maternelle sont-elles prédictives des difficultés de lecture au cours préparatoire ? *Approche Neuropsychologique des Apprentissages chez l'Enfant, 91*, 5-17.

Guimard, Ph., Cosnefroy, O., & Florin, A. (2007). Evaluation des comportements et des compétences scolaires par les enseignants et prédiction des performances et des parcours à l'école élémentaire et au collège. *L'Orientation Scolaire et Professionnelle, 36*, 179-202.

Guimard, Ph., & Florin, A. (2008). Evaluer les compétences et les trajectoires scolaires : nouvelles perspectives d'évaluation. In J. M. Hoc & Y. Corson (Eds.), *Actes du Congrès 2007 de la Société Française de Psychologie* (pp. 141-149).

Guttentag, R. E., & Haith, M. M. (1978). Automatic processing as a function of age and reading ability. *Child Development, 49*, 707-716.

Hamzaoui, H. (1988). Retard scolaire, langage et inégalité sociale. *Cahier du C.E.R.E.S, série psychologique, 5*, 193-278.

Holm, A., & Dodd, B. (1999). A longitudinal study of the phonological development of two Cantonese-English bilingual children. *Applied Psycholinguistics*, *20*, 349-376.

Huteau, M. (1995). *Manuel de psychologie différentielle*. Dunod, Paris.

Ibrahim, R., Eviatar, Z., & Aharon-Peretz. J. (2002). The characteristics of Arabic orthography slow its processing. *Neuropsychology*, *3*, 322, 326.

Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2007). Metalinguistic Awareness and Reading Performance: A Cross Language Comparison. *Journal of Psycholinguistic Research*, *36*, (4), 297-317.

INSERM. (2007). *Dyslexie, dysorthographie, dyscalculie: Bilan des données scientifiques*. Paris, Editions INSERM.

Jimenez, J. E., Siegel, L. S., & Lopez, M. R. (2003). The relationship between IQ and reading disabilities in English-speaking Canadian and Spanish children. *Journal of Learning Disabilities, 36 (1)*, 15-23.

Jorm, A. F., Share, D. L., MacLean, R., & Matthiews, R. G. (1984). Phonological recoding skill and learning to read : A longitudinal study. *Applied Psycholinguistics, 5*, 201-207.

Juel C., Griffith P. L., & Gough P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*, 243-255.

Kayser, D., & Houdé, O. (1998). *Vocabulaire de sciences cognitives*. PUF.

Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. V. (2008). Longitudinal predictors of word reading development. *Canadian Psychology, 2*, 103-110.

Kirby, J. R. Parrila, R. K., & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology, 3*, 453-464.

Khomsi, A. (1990). *Epreuve collective d'évaluation de la compétence en lecture : E20*. Paris, Editions du CPA.

Khomsi, A. (1994). *Epreuve d'Identification de Mots Ecrits : IME*. Paris, Editions du CPA.

Kolinsky. R., Morais, J., & Cluytens, M. (1995). Intermediate representations in spoken word recognition: Evidence from word illusions. *Journal of Memory and Language*, *34*, 19-40.

Labécd. (2001). *Présentation des épreuves pour la GSM*. Université de Nantes.

Laxon, V., Richard, M., & Coltheart, V. (1992). Children read affixed words and non-words. *British Journal of psychology, 83*, 407-423.

Lecoq, P. (1991). *Apprentissage de la lecture et dyslexie*. Bruxelles : Mardaga.

Lecoq, P., Leuwers, C., Casalis, S., & Watteau, N. (1996). *Apprentissage de la lecture et compréhension d'énoncés.* Paris : Presse Universitaire du Septentrion.

Lecomte, G. (1968). *Grammaire de l'arabe*. PUF.

Leong, C. K., Hau, K. T., Cheng, P. W., & Tan, L.H. (2005). Exploring two-wave reciprocal structural relations among orthographic knowledge, phonological sensitivity, and reading and spelling of English words by Chinese students. *Journal of Educational Psychology, 97*, 591-600.

Leybeart, J., & Alegria, J. (1995). Spelling development in deaf and hearing children : Evidence for use of morphophonological regularities in Frensh, *Reading and Writing : An Interdisciplinary Journal, 7*, 1-21.

Leybaert, J., Alegria, J., Deltour, J. J., & Skinkel, R. (1994). Apprendre à lire : rôle du langage, de la conscience phonologique et de l'école. In J. Grégoire & B. Piérart (Eds.),

Evaluer les troubles de la lecture : les nouveaux modèles théoriques et leurs implications diagnostiques (pp. 147-171). Bruxelles, De Boek Université.

Liberman, I. Y., & Shankweiler, D. (1989). Phonologie et apprentissage de la lecture : une introduction. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques* (pp. 23-42). Paris, Neuchâtel, Delachaux et Niestlé.

Liberman, I. Y., Shankweiler, D., Fisher, M., & Carter, B. (1974). Explicit syllabe and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18,* 201-212.

Lieury, A., & De La Haye, P. (2004). *Psychologie cognitive de l'éducation*. Dunod.

Lindsey, K. A., Manis, F. R., & Baily, C. E. (2003). Prediction of first-grade reading in spanish-speaking English-language learners. *Journal of Educational Psychology, 3,* 482-494.

Lundberg, I. (1982). Linguistic awareness as related to dyslexia, in Y. Zotterman (Ed). *Dyslexia: Neuronal, Cognitive and Linguistics aspects* (pp. 141-153). New York, Pergamon.

Lyster, S. A. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing: An Interdisciplinary Journal, 15,* 261-294.

Mann, V. A. (1989). Les habiletés phonologiques: prédicteurs valides des futures capacités en lecture. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques* (pp. 221-240). Neuchâtel, Delachaux et Niestlé.

Marec-Breton, N. (2009). Quelles informations morphologiques les apprentis lecteurs utilisent-ils au cours de leur lecture ? In N. Marec-Breton A-S. Besse, F. de la Haye, N. Bonneton & E. Bonjour (Eds), *L'apprentissage de la langue écrite : Approche cognitive* (pp. 81-92). Rennes : PUR.

Marec-Breton, N., Gombert, J.-E., & Colé, P. (2005). Traitements morphologiques lors de la reconnaissance des mots écrits chez des apprentis lecteurs, *L'année Psychologique, 105*, 9-45.

Marec-Breton, N. (2003). *Les traitements morphologiques dans l'apprentissage de la lecture.* Thèse de doctorat en psychologie non publiée, Université Rennes 2.

Marivain, Th. (1992). Les déterminants cognitifs de l'apprentissage de la lecture : une revue critique des comparaisons entre bons et mauvais lecteurs. *L'Année Psychologique, 92,* 559-582.

Masonheimer, P. E., Drum, P. A., & Ehri, L. C. (1984). Does environmental print identification lead children into word reading? *Journal of Reading Behaviour, 16,* 257-271.

Mc Conkie, G. W., & Zola, D. (1981). Language constraints and the functional stimulus in reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive process in reading* (pp. 155-175). Hillsdale, NJ: Lawrence Erlbaum.

Meljac, C., Kugler, M., & Mogenet, J. l. (2001). Prevision, prediction, anticipation, et apprentissage de la lecture. *Pratiques psychologiques, 1,* 81-97.

Ministère de l'Education et de la Formation tunisien (2004). *Rapport national sur le développement de l'éducation 2000-2004.*

Ministère de l'Education et de la Formation tunisien (2007). *Plan d'action national de l'éducation pour tous : évaluation à mi-parcours*. Commission nationale tunisienne pour l'éducation, la science et la culture.

Ministère de l'Education Nationale français (2001). *Evaluation à l'école primaire. Evaluation et aide aux apprentissages en grande section de maternelle et en cours préparatoire : identification des compétences et repérage des difficultés des élèves*. Direction de la programmation et du développement, mission de l'évaluation.

Morais, J. (1991). Constraints on the development of phonemic awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *phonological process in literacy. A tribute to Isabelle Y. Liberman* (pp. 5-27). Hillsdale, NJ: Lawrence Erlbaum.

Morais, J. (1994). *L'art de lire*. Paris : Editions Odile Jacob.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7,* 323-331.

Mousty, Ph., Leybeart, J., Alégria, J., Content, A., & Morais, J. (1994). BELEC : une batterie d'évaluation du langage écrit et de ses troubles. In J. Grégoire & B. Piérart (Eds.), *Evaluer les troubles de la lecture : les nouveaux modèles théoriques et leurs implications diagnostiques* (pp. 127-145). Bruxelles, De Boek Université.

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.

Muter.V., Hulme, C., Snowling, M., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental psychology, 5*, 665-681.

Naglieri, J. A. (1998). *NNAT : Test d'Aptitudes Non Verbales de Naglieri (forme A)*. Paris : ECPA.

Nicholson, T. (1991). Do children read word better in context or in list ? A classic study revisited. *Journal of Educational psychology, 83*, 444-450.

Nocus, I., & Gombert, J.-E. (1997). Conscience morpho-syntaxique et apprentissage de la lecture. *Revue de Psychologie de l'Education, 2*, 71-102.

Peerman, R. (1991). La médiation phonologique dans la reconnaissance des mots écrits. In R. Kolinsky, J. Morais & J. Segui (Eds.), *La reconnaissance des mots dans les différentes modalités sensorielles : études de psycholinguistique cognitive* (pp. 119-163). Paris, PUF.

Peerman, R. (1992). Lecture, écritures, orthographes. In P. Lecoq (Ed.). *La lecture : processus, apprentissage, troubles* (pp. 55-77). Presses Universitaires de Lilles.

Perfetti, C. A. (1989). Representations et prise de conscience au cours de l'apprentissage de la lecture. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur: recherches empiriques et implication pédagogiques* (pp. 61-84). Neuchâtel, Delachaux et Niestlé.

Perfetti, C. A., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly, 33*, 283-319.

Perfetti, C. A., Goldman, S. R., & Hogaboam, T. W. (1979). Reading skill and the identification of words in discourse content. *Memory & Cognition, 7*, 273-282.

Peterson, D. K., & Elbro, C. (1999). Pre-school prediction and prevention of dyslexia: A longitudinal study with children of dyslexic parents. In T. Nunes (Ed.), *learning to read: An integrated view from research and practice* (pp. 133-154). Dordrecht: Kluwer.

Piquard-Kipffer, A. (2003). *Prédiction de la réussite ou de l'échec spécifiques en lecture au cycle 2 (suivi d'une population contrôle de la moyenne section de maternelle à la 2ème année de scolarisation primaire).* Thèse de Doctorat non publiée, Université Paris 7 Denis Diderot.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879-891.

Quay, L.C., & Steele, D. C. (1998). Predictive children achievement from teacher judgements : an alternative to standardized testing. *Early Education and Development, 9 (3),* 207-218.

Ravid, D. (1996). Accessing the mental lexicon: Evidence from incompatibility between representation of spoken and written Hebrew. *Linguistics, 3,* 1219-1246.

Ravid, D. (2001). Learning to spell in Hebrew : phonological and morphological factors. *Reading and writing: An Interdisciplinary Journal, 14*, 459-485.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sound depends on knowing alphabetic reading, *cognition, 24*, 31-44.

Reuchlin, M. (1990). *Cognition: L'individuel et l'universel*. Paris, PUF.

Reuchlin, M. (1991). *Les différences individuelles à l'école*. Paris, PUF.

Rohl, M., & Pratt, C. (1995). Phonological awareness, verbal working memory and the acquisition of literacy. *Reading and Writing: An Interdisciplinary Journal, 7,* 327-360.

Roman, G. (1982). *Approche expérimentale de la lecture de l'arabe*. Thèse de doctorat de 3ème cycle en psychologie non publiée, Université de Provence, Aix en Provence.

Roman, A. (1983). La genèse de la langue arabe et la constitution de ses formes verbales. *Cahiers du C.E.R.E.S, 5,* 29-45.

Roman, A., Pavard, B., & Asselah, B. (1985). Traitement perceptif des phrases ambigües en arabe. *Cahiers de Psychologie Cognitive, 5(1)*, 5-22.

Romdhane, M. N. (2003). Déterminants de l'acquisition de la lecture compréhension en arabe. In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 99-114). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Roussel, P., Durrieu, F., Compoy, E., & El Akremi, A. (2002). *Méthodes d'équations structurelles : recherches et application en gestion*. Edition Economica.

Roussel, P., & Wacheux, F. (2006). *Management des ressources humaines : méthodes de recherche en sciences humaines et sociales*. De Boek University.

Sabatier, C., Baudet, V., Habib, M., Prost, V., Thomas, C., & Rey, V. (2003). Déficits phonologiques et morphologiques chez des enfants dyslexiques francophones. In M. N. Romdhane, J.-E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 179-193). Presses Universitaires de Rennes & Centre de Publication Universitaire de Tunis.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of arabic diglossia. *Applied Psycholinguistics, 24(3)*, 431-451.

Saiegh-Haddad, E. (2004). The impact of phonemic and lexical distance on the phonological analysis of words and pseudowords in a diglossic context. *Applied Psycholinguistics, 25*, 495-512.

Scarborough, H. S. (1989). Prediction of reading disability from familial and individual differences. *Journal of Educational Psychology, 1*, 101-108.

Scarborough, H. S. (1998). Predicting the future achievement of second graders with reading disabilities: contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia, 48*, 115-136.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Seymour, P. H. K. (1986). *A cognitive analysis of dyslexia*. London: Routledge.

Seymour, P. H. K., & Elder, L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology, 3*, 1-37.

Seymour, P. H. K., & Evans, H. M. (1992). Beginning reading without semantics: A cognitive study of hyperlexia. *Cognitive Neuropsychology, 2*, 89-129.

Shankweiler, D., Crain, S., Brady, S., & Maracuso, P. (1992). Identifying the cause of reading disability. In P. B. Gough, L. C. Ehri & R. Treinman (Eds.), *Reading Acquisition* (pp. 275-305). Hillsdale: New Jersey, Lawrence Erlbaum Associates.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218.

Schatschneider, Ch., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: a longitudinal comparative analysis. *Journal of Educational Psychology, 2*, 265-282.

Siegel, L. S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. *Canadian Journal of Psychology, 42*, 201-215.

Siegel, L. S. (1989). Why we do not need intelligence test scores in the definition and analyses of learning disabilities. *Journal of Learning Disabilities, 22*, 514-518.

Siegel, L. S. (1990). IQ and learning disabilities: R.I.P. In H. L. Swanson & B. K. Keogh (Eds.), *Learning disabilities: Theoretical and research issues*. Hillsdale : Elbraum.

Siegel, L. S. (1993). Phonological processing in reading as the basis of reading disabilities. *Developmental Review, 13*, 246-257.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973-980.

Snoussi, F. (2007). Effets de la situation de diglossie sur l'apprentissage de la lecture en langue arabe. *Traverse, 6*, 63-77.

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: the effects of oral language, emergent literacy, and behavior skills. *Developpemental Psychology, 1*, 225-234.

Sprenger-Charolles, L. (1992). L'évolution des mécanismes d'identification des mots. In M. Fayol, J. E. Gombert, L. Sprenger-Charolles & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 141-193). Paris, PUF.

Sprenger-Charolles, L., & Khomsi, A. (1989). Les stratégies d'identification des mots écrits dans un contexte-image : comparaison entre « bons » et « mauvais » lecteurs. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques* (pp. 307-330). Neuchâtel, Delachaux et Niestlé.

Sprenger-Charolles, L., & Casalis, S. (1996). *Lire. Lecture et écriture : acquisition et troubles du développement*. Paris, PUF.

Sprenger-Charolles, L., & Colé, P. (2003). *Lecture et dyslexie : approche cognitive*. Paris, Dunod.

Sprenger-Charolle, L., Colé, P., Lacert, P., & Serniclaes, W. (2000). On subtypes of developmental dyslexia: evidence from processing time and accuracy scores. *Canadian Journal of Experimental Psychology, 54*, 88-104.

Stanovich, K. E. (1979). Attentional and automatic contexte effects in reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive process in reading* (pp. 241-267). Hillsdale, Laurence, Erlbaum.

Stanovich, K. E. (1986). Matthew effects in reading: some consequences of individual differences in acquisition of literacy. *Reading research Quarterly, 4*, 360-406.

Stanovich, K. E. (1989). L'évolution des modèles de lecture et de l'apprentissage de la lecture. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques* (pp. 43-60). Neuchâtel, Delachaux et Niestlé.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. (1984). Assessing phonological awarness in kindergarten: issues of task comparability. *Journal of Experimental Child Psychology, 38*, 175-190.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology, 6*, 934-947.

Swanson, H. L. (1999). Reading comprehension and working memory in learning-disabled readers: is the phonological loop more important than the executive system. *Journal of Experimental Child Psychology 72*, 1–31.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14*, 638-647.

Taouk, M., & Coltheart, M. (2004). The cognitive process involved in learning to read in Arabic. *Reading and Writing: An Interdisciplinary Journal, 17*, 27-57.

Tornéus, M. (1990). The importance of metaphonological and metamorphological habilities for different phases of reading development. In G. T. Pavlidis (Ed.). *Perspectives on Dyslexia* (vol. 2, pp. 153-173). New York: Jorm Wiley & Son.

Tourette, C. (2006). *Evaluation psychologique des enfants avec déficiences ou troubles du comportement*. Dunod.

Treiman, R. (2006). Knowledge about letters as a foundation for reading and spelling. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 581–599). Mahwah, NJ: Erlbaum.

Treiman, R. (1994). Acquisition de l'écriture en anglais. In J. P. Jaffré, L. Sprenger-Charolles & M. Fayol (Eds.), *Les actes de la villette. Lecture-écriture : acquisition* (pp. 124-136). Paris, Nathan.

Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P. Gough, L. Ehri, & R. Treinman (Eds.), *Reading Acquisition* (pp. 273-281). Hillsdale, NJ : Erlbraum.

Treiman, R. (1989). Le rôle des unités intra-syllabiques dans l'apprentissage de la lecture. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : Recherches empiriques et implications pédagogiques* (pp. 241-259). Neuchâtel, Delachaux et Niestlé.

Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology, 39*, 161-181.

Tunmer, W. E. (1989). Conscience phonologique et acquisition de la langue écrite. In L. Rieben & C. A. Perfetti (Eds.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques* (pp. 197-220). Neuchâtel, Delachaux et Niestlé.

Valdois, S., Bosse, M .L., & Tainturier, M. J. (2004). Cognitive correlates of developmental dyslexia : Review of evidence for a selective visual attentional deficit. *Dyslexia*, *10*, 1-25.

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

Van Orden, G. (1987). A rows is a rose: spelling, sound and reading. *Memory and Cognition, 5*, 181-198.

Vellutino, F. R., & Scanlon, D. M. (2002). The interactive strategies approach to reading intervention. *Contemporary Educational Psychology*, 27, 573-635

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role on the acquisition of reading skills. *Psychological Bulletin, 2*, 192-212.

Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Children's and adult's use of spelling sound information in three reading tasks. *Memory and Cognition, 12*, 293-305.

Wimmer, H. (1994). L'acquisition de la lecture dans une orthographe plus régulière que celle de l'anglais: points de divergence. In J. P. Jaffré, L. Sprenger-Charolles & M. Fayol (Eds.), *Les actes de la villette. Lecture-écriture: acquisition* (pp. 97-106). Paris, Nathan.

Wimmer, H., & Hammer, P. (1990). How German speaking first graders read and spell : doubts on the importance of the logographic stage. *Applied Psycholinguistics, 11*, 349-368.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development : word recognition in English and German children. *Cognition, 51,* 91-103.

Wolf, M., & Bowers, P. G. (1999). The double deficit hypothesis for the developmental dyslexia. *Journal of Educational Psychology, 91* (3), 415-438.

Zagar, D. (1992). L'approche cognitive de la lecture : de l'accès au lexique au calcul syntaxique. In M. Fayol, J. E. Gombert, P. Lecoq, L. Sprenger-Charolle & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 15-72). Paris, PUF.

Zazzo, B. (1978). *Un grand passage : de l'école maternelle à l'école élémentaire*. Paris, PUF.

Zazzo, R. (1967). *Manuel pour l'examen psychologique de l'enfant*. Neuchâtel, Delachaux & Niestlé.

Zorman, M. (1999). Evaluation de la conscience phonologique et entraînement des capacities phonologiques en grande section de maternelle. *Rééducation Orthophonique, 197,* 139-157.

Zorman, M., Lequette, Ch., & Pouget, G. (2004). Dyslexies : intérêt d'un dépistage et d'une prise en charge précoce de l'école. Evaluation du BSEDS 5-6. In M.-N. Metz-Lutz, E. Demont, C. Seegmuller, M. De Agostini & N. Bruneau (Eds.). *Développement cognitif et troubles des apprentissages : évaluer, comprendre, rééduquer et prendre en charge* (pp. 245-270). Solal Editeur, Marseille.

Liste des tableaux

Tableau 1: Exemple en arabe d'une forme non vocalisée à laquelle peuvent correspondre trois mots de significations différentes..... 40
Tableau 2 : Liste de l'alphabet arabe..... 46
Tableau 3 : Exemple de dérivation à partir de la racine k-t-b..... 50
Tableau 4 : Exemples de diglossie..... 52
Tableau 5 : Caractéristiques de la population..... 152
Tableau 6 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus aux épreuves langagières à T1..... 163
Tableau 7 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus aux épreuves cognitives à T1..... 164
Tableau 8 : Performances moyennes de réussite, écart-types, minima, maxima obtenus aux épreuves langagières à T2..... 165
Tableau 9 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus à l'épreuve de lecture IME1..... 166
Tableau 10 : Performances moyennes de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus aux épreuves de lecture IME2 et Timé2..... 167
Tableau 11 : Performance moyenne de réussite (pourcentage moyen de réussite), écart-types, minima, maxima obtenus à l'épreuve de compréhension à T3..... 168
Tableau 12 : Matrice des corrélations entre les compétences langagières et cognitives évaluées en CP..... 169
Tableau 13 : corrélations entre les différentes mesures à T2..... 170
Tableau 14 : Corrélations entre les différentes mesures à T1 et T2..... 171
Tableau 15 : Corrélations entre les différentes épreuves de lecture à T1 et T2..... 172
Tableau 16 : Matrice de l'analyse factorielle en composantes principales des mesures à T1 (les saturations inférieures à 0.5 ne sont pas indiquées)..... 173
Tableau 17 : Corrélations entre les différents facteurs à T1..... 174
Tableau 18: Matrice de l'analyse factorielle en composantes principales des mesures à T2 174
Tableau 19: Matrice de l'analyse factorielle en composantes principales des mesures à T3 175
Tableau 20 : Corrélations entre les compétences cognitives et langagières en CP et les performances en lecture en 1ère année (IME1 : identification de mots écrits à T2)..... 178
Tableau 21 : Corrélations entre les compétences à T2et les performances en lecture en 1ère année (IME1 : identification de mots écrits à T2)..... 178
Tableau 22: Analyse de régression pas à pas entre les mesures à T1 et les performances en identification de mots à T2..... 179
Tableau 23 : Analyse de régression pas à pas entre les mesures à T2 et les performances en identification de mots à T2..... 179
Tableau 24 : Corrélations entre les mesures à T0 et les performances en identification de mots à T2..... 180
Tableau 25 : Corrélations entre les mesures à T2et les performances en identification de mots à T3..... 181
Tableau 26 : Corrélations entre les mesures à T1et les performances en compréhension écrite à T3..... 182
Tableau 27 : Corrélations entre les mesures à T2 et les performances en lecture-compréhension à T3..... 182
Tableau 28 : Analyse de régression pas à pas entre les mesures à T1 et les performances en identification de mots à T3..... 183

Tableau 29 : Analyse de régression pas à pas entre les mesures à T2 et les performances en identification de mots à T3 183
Tableau 30 : Analyse de régression pas à pas entre les mesures à T1 et les performances à l'épreuve Timé2 à T3 184
Tableau 31 : Analyse de régression entre les mesures à T2 et les performances à l'épreuve timé2 à T3 184
Tableau 32 : Analyse de régression entre les mesures à T1 et les performances à l'épreuve de lecture-compréhension à T3 185
Tableau 33 : Analyse de régression entre les mesures à T2 et les performances à l'épreuve de lecture-compréhension à T3 185
Tableau 34 : Récapitulatif des meilleurs prédicteurs de l'apprentissage de la lecture en CP (T1) 187
Tableau 35 : Récapitulatif des meilleurs prédicteurs de l'apprentissage de la lecture en 1ère année (T2) 187
Tableau 36 : Estimation de la validité convergente du modèle de mesure 203
Tableau 37 : Corrélation entre les variables latentes 204
Tableau 38: Estimation de la variance expliquée 205
Tableau 39 : Indices d'ajustement du modèle de mesure 207
Tableau 40 : Indices d'ajustement du modèle respécifié 207
Tableau 41: Test de significativité des paramètres du modèle des prédicteurs de l'apprentissage de la lecture en arabe 209
Tableau 42: Estimation des effets directs et indirects des connaissances phonologiques sur les performances en lecture 210
Tableau 43 : Indices d'ajustement du modèle des prédicteurs de l'apprentissage de la lecture en arabe 211
Tableau 44 : Matrice des corrélations entre les items du questionnaire 217
Tableau 45 : Statistique descriptive du questionnaire 218
Tableau 46: Valeurs du coefficient Alpha en retirant un item 218
Tableau 47 : Test de validité convergente du construit « compétences langagières » 219
Tableau 48 : Test de validité convergente du construit « comportement en classe » 220
Tableau 49: Test de validité convergente du construit « compétences scolaire » 220
Tableau 50 : Corrélations entre les variables latentes 221
Tableau 51 : Indices d'ajustement du modèle 222
Tableau 52 : Corrélations entre les items du questionnaire à T1 et les performances en identification de mots à T2 (IME1) 223
Tableau 53 : Analyse de régression pas à pas entre les items du questionnaire à T1 et les performances en identification de mots à T2 (IME1) 224
Tableau 54 : Corrélations entre les items du questionnaire aux enseignants et les performances en lecture à T3 226
Tableau 55 : Analyses de régression pas à pas entre les items du questionnaire et les performances à l'épreuve IME2 à T3 227
Tableau 56 : Analyses de régression pas à pas entre les items du questionnaire et les performances au Timé2 à T3 227
Tableau 57 : Analyse de régression pas à pas entre les scores globaux au questionnaire et aux mesures langagières à T1 et les scores en lecture à T3 228
Tableau 58 : Analyse de régression pas à pas entre les items du questionnaire et les performances en lecture-compréhension à T3 228
Tableau 59 : Ajustement et pouvoirs explicatifs du modèle 230
Tableau 60 : Estimation par le maximum de vraisemblance de la significativité des paramètres du modèle 231

Tableau 61 : Récapitulatif des items du questionnaire les plus prédictifs des performances en lecture à T2 et T3 234
Tableau 62 : Moyennes, écart-types et centiles 25 des performances aux scores globaux langagiers et cognitifs et au questionnaire enseignants 238
Tableau 63 : Enfants prédits en difficulté et en réussite en fonction de leurs scores globaux en langage, dans le domaine cognitif et au questionnaire enseignant 238
Tableau 64 : Moyennes, écart-types et centiles 25 des performances aux épreuves de lecture à T2 et T3 239
Tableau 65 : Enfants prédits en difficulté et en réussite en fonction de leurs scores aux épreuves de lecture 239
Tableau 66 : Calcul des indices et valeurs prédictives 240
Tableau 67 : Valeur prédictive individuelle des scores langagier et cognitif et du questionnaire enseignants à T1 sur les performances à l'épreuve IME1 241
Tableau 68 : Valeur prédictive individuelle des scores langagier et cognitif et du questionnaire enseignants sur les performances à l'épreuve IME2à T3 242
Tableau 69 : Valeur prédictive individuelle des des scores langagier et cognitif et du questionnaire enseignants sur les performances à l'épreuve Timé2 à T3 243
Tableau 70 : Valeur prédictive individuelle des scores langagier et cognitif et du questionnaire enseignants sur les performances en lecture-compréhension à T3 244
Tableau 71: Différences significatives sur les compétences langagières à T1 selon les niveaux de performances aux épreuves d'identification de mots à T2 (faibles=FL ; moyens=ML ; bons=BL) 246
Tableau 72 : Différences significatives sur les compétences cognitives à T1 selon les niveaux de performances aux épreuves d'identification de mots à T3 (faibles=FL ; moyens=ML ; bons=BL) 248
Tableau 73 : Différences significatives sur les épreuves passées à T2 selon les niveaux de performances aux épreuves d'identification de mots à T3 (faibles=FL ; moyens=ML ; bons=BL) 249

Liste des figures

Figure 1 : Modèle connexionniste adapté de Seidenberg & McClelland (1989) 36
Figure 2 : Apprentissage implicite et explicite de la lecture (d'après Gombert, 2003) 37
Figure 3 : Modèle de lecture chez le lecteur habile en arabe (d'après Abu-Rabia, 1997) 61
Figure 4 : Les différents niveaux de segmentation phonologique (Gombert & al., 2000)..... 66
Figure 5 : Le développement métalinguistique (Gombert & al., 2004) 71
Figure 6 : Modèle théorique proposant une intervention précoce de la dimension morphologique dans l'apprentissage de la lecture (d'après Colé & al., 2004) 81
Figure 7 : Scores moyens (en %) obtenus aux différentes épreuves langagières à T1 165
Figure 8 : performances exprimées en pourcentages moyens de réussite obtenues aux différents sous-scores des épreuves d'identification de mots à T3 (IME2 et Timé2)..... 167
Figure 9 : Moyenne des corrélations obtenues avec les épreuves de lecture en 1ère année (T2) et en 2ème année (T3) 186
Figure 10 : Modèle théorique des prédicteurs de l'apprentissage de la lecture en langue arabe 201
Figure 11 : Modèle d'équations structurelles des prédicteurs de l'apprentissage de la lecture en langue arabe..... 208
Figure 12 : Résultats de l'analyse factorielle confirmatoire 221
Figure 13 : Modèle d'équations structurelles des relationsprédictives entre les dimensions du questionnaire et les performances en lecture à T3 229
Figure 14 : Performances obtenues par les trois groupes de lecteurs aux différentes épreuves langagières et au questionnaire enseignants en CP 245
Figure 15 : Performances obtenues par les trois groupes de lecteurs de 2ème année aux différentes épreuves cognitives en CP..... 247
Figure 16 : Performances obtenues par les trois groupes de lecteurs aux différentes épreuves passées en 1ère année..... 248

ANNEXES

Annexe I : Les épreuves à la fin de la classe préparatoire (T1)

En mai-juin 2005, nous avons administré aux enfants de classe préparatoire différentes épreuves d'évaluation du langage oral et écrit et des compétences cognitives lors de passations collectives et individuelles. Compte tenu de la situation de diglossie, toutes les consignes sont données en arabe dialectal. Avant chaque épreuve, un feed-back correctif est fourni en cas de réponse erronée.

Epreuves d'évaluation des connaissances phonologiques

Tâche de détection d'intrus de rime

Cette tâche vise à évaluer la sensibilité des jeunes enfants à la rime.

Passation : collective

Consigne :

Regardez ces images. Je vais vous donner chaque fois trois mots et vous allez entourer l'image du mot qui ne finit pas pareil que les deux autres. Je vais vous expliquer encore plus par des exemples.

شوفو التصاور هذوما ، آنا في كل مرّة بش نقلكم ثلاثة كلمات و انتوما بش ادوّرو التصويرة إلى اسمها ما يوفاش كيما لوخرين – وهاو بش نجربو مع بعضنا

شوفو فمّا : كار- ساق - فار

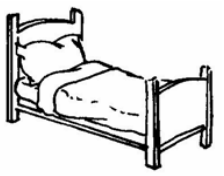
أنا هي الكلمة إلّى ما تبداش كما لوخرين

Matériel :

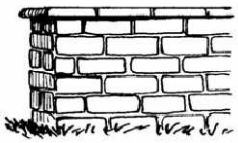
كلب	قلب	فرش
خيط	عين	حيط
خشم	جبل	حبل
كاس	نار	فاس
قجر	حجر	مقص
قلم	تبن	جبن



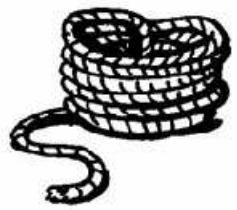


























Tâche de détection de l'intrus de syllabe initiale

Passation : collective

Consigne :

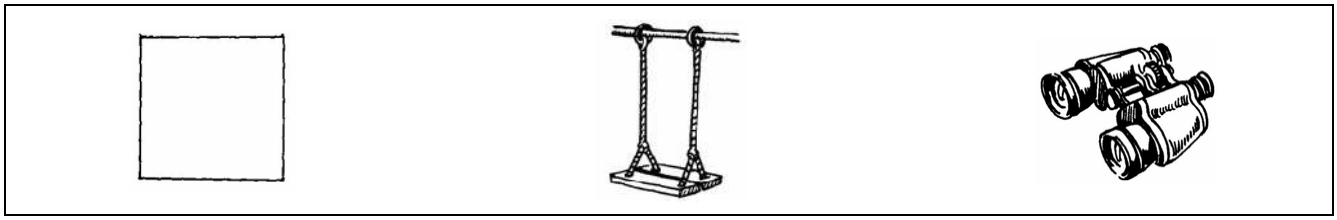
Regardez ces images. Je vais vous donner chaque fois trois mots et vous allez entourer l'image du mot qui ne commence pas pareil que les deux autres. Je vais vous expliquer encore plus par des exemples.

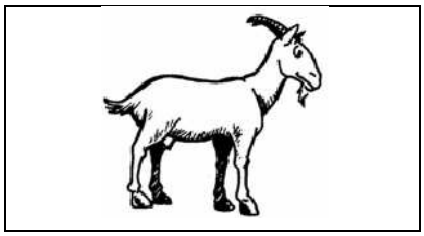
شوفو التصاور هذوما ، آنا في كل مرّة بش نقلكم ثلاثة كلمات و انتوما بش ادوّرو التصويرة إلى اسمها ما يبداش كيما لوخرين – وهاو بش نجربو مع بعضنا

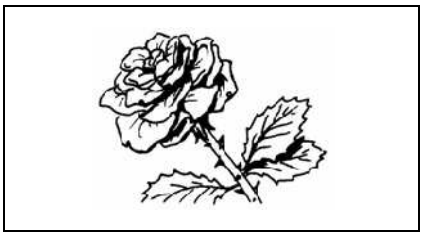
مكبرة درجيحة مُربّع

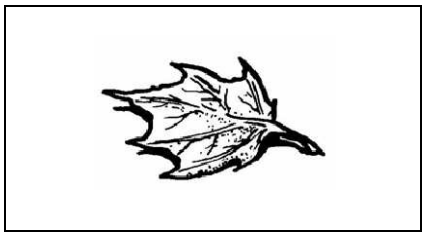
Matériel :

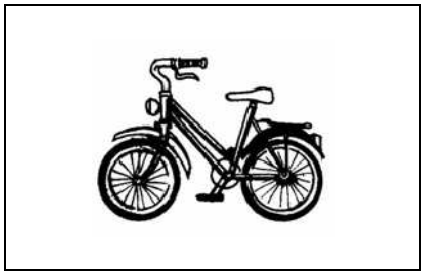
ورقة	وردة	معزة
مفتاح	بسباس	بسكلات
سبعة	قهوة	سبتة
كورة	كوبة	دودة
حلزون	أمبوبة	أمبولة
بانو	شيتة	بالة

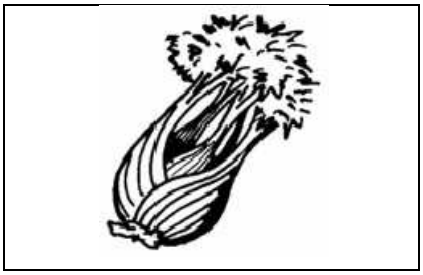




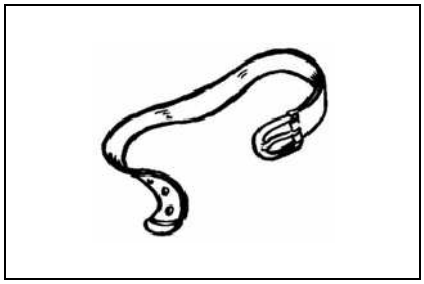


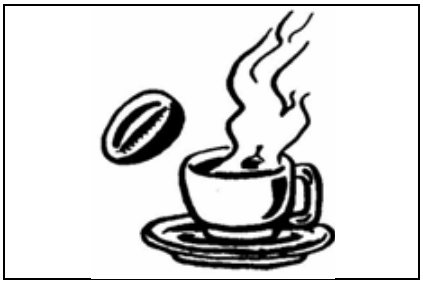


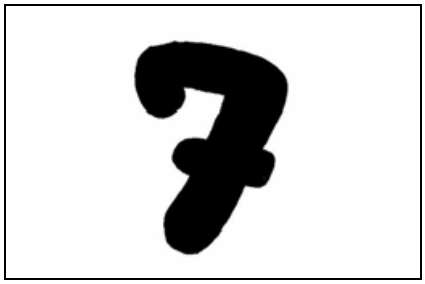


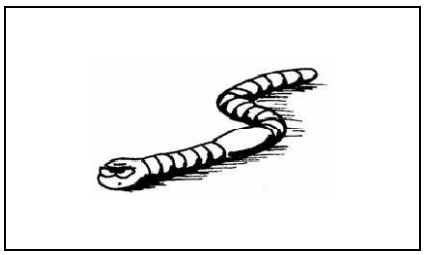


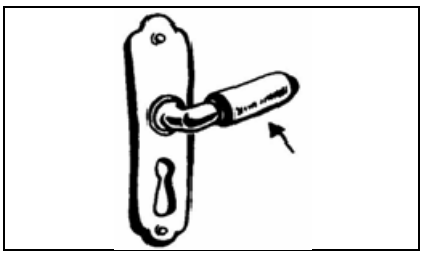


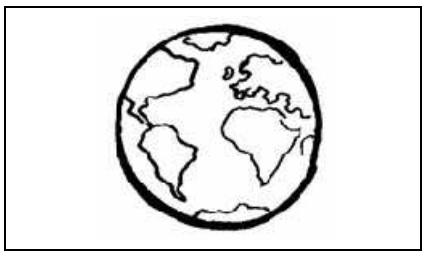


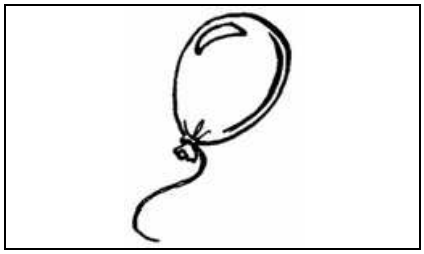


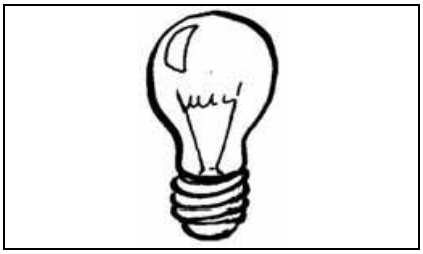


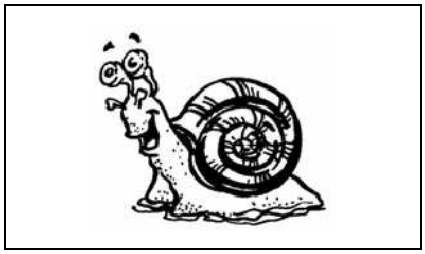


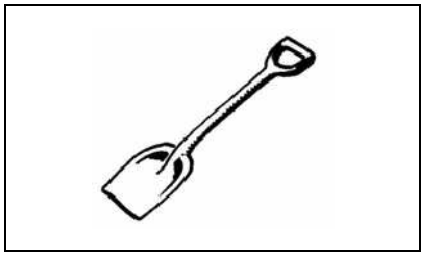


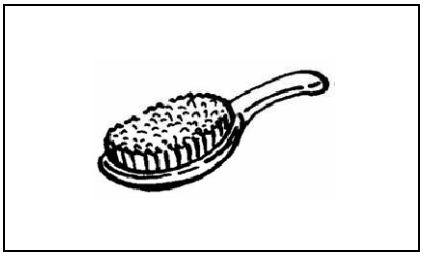


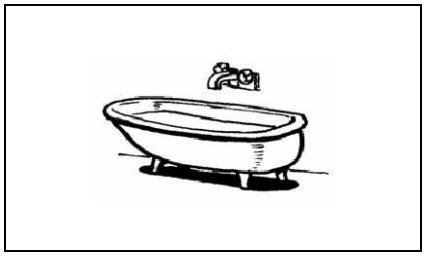












Tâche de comptage et segmentation syllabique

Passation : Individuelle

Consigne :

Je vais te lire des mots en arabe dialectal. Pour chaque mot, tu vas compter, à l'aide de tes doigts, le nombre de morceaux qu'il comporte. Ensuite, tu vas nommer ces morceaux séparément et en ordre. Je vais t'expliquer encore plus par des exemples.

بش نقلّك كلمات بالدّارجة، كل كلمة بش تحسب بصوابعك و تقلّي قدّاش فيها من جزء ، امبعد بش تسمّيهملي منظمين بالواحد بالواحد– وهاو بش نجربو مع بعضنا

Matériel :

كرَاسة [korrasa]	دودة [du:da]	درجيحة [durdziħa]	
بحر [bħar]	حلزون [ħalazu:n]	فيل [fi:l]	
تصويرة [tasˁwira]	ثوم [θu:m]	يِجري [jiʤri]	
قلم [qlam]	أحمر [aħmar]	صابون [sa:bun]	

Tâche de suppression syllabique

Passation : Individuelle

Consigne :

Je vais te lire des mots en arabe dialectal. Chaque fois, je vais enlever un morceau et tu dois me dire ce qui reste du mot. Je vais t'expliquer encore plus par des exemples.

بش نقلّك كلمات بالدّارجة، في كل مرّة انحي منهم جزء وإنت تقلّي شنوّة قعد في الكلمة– وهاو بش نجربو مع بعضنا

Matériel :

كوجينة	[ku:dzi:na]	شجرة	[ʃodzra]	شكلاطة	[ʃokla:ta]
كانون	[ka:nu:n]	نمّالة	[nimma:la]	فليجة	[fali:dza]
موزة	[mu:za]	واحد	[wa:ħid]	سروال	[sirwa:l]
طربوشة	[tarbuʃa]	مفتاح	[mifta:ħ]	جلبانة	[dzilba:na]

Epreuves d'évaluation des connaissances phonémiques

Tâche de détection de l'intrus phonémique

Passation : collective

Consigne :

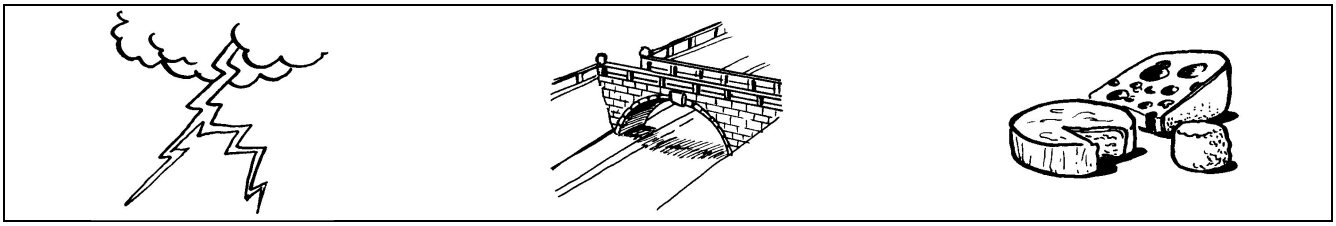
Regardez ces images. Je vais vous donner chaque fois trois mots et vous allez entourer l'image qui correspond au mot qui ne commence pas par le même son que les deux autres. Je vais vous expliquer encore plus par des exemples.

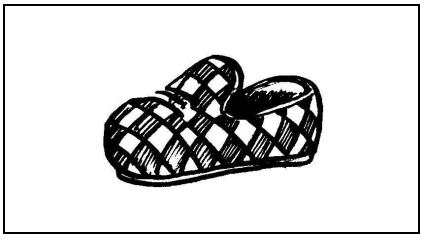
شوفو التصاور هذوما ، آنا في كل مرّة بش نقلكم ثلاثة كلمات و انتوما بش ادوّرو التصويرة إلى اسمها ما يبداش بنفس الصوت هي و لخرين – وهاو بش نجربو مع بعضنا

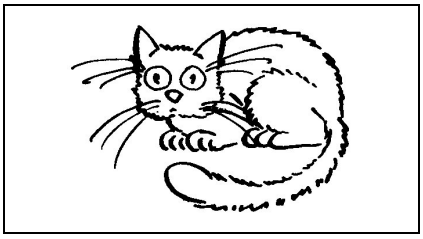
جُبن - جِسر - بَرق

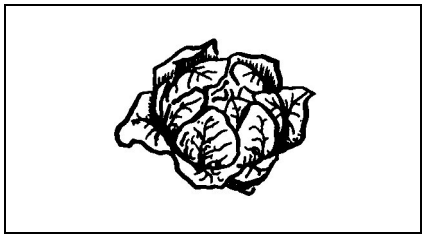
Matériel :

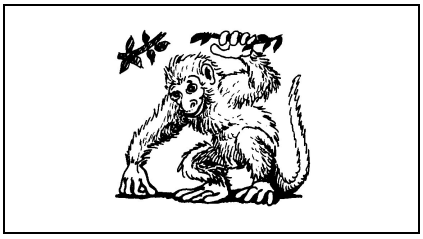
خَس	**قِط**	**خُف**
خبز	**قَلب**	**قِرد**
مَكتوب	**مِنديل**	**قَرصان**
صُمعة	**كِسوة**	**صَفحة**
مُشط	**رَأس**	**مِلح**
عُلبة	**عَنزة**	**طِفلة**

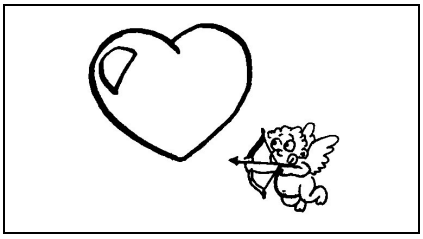


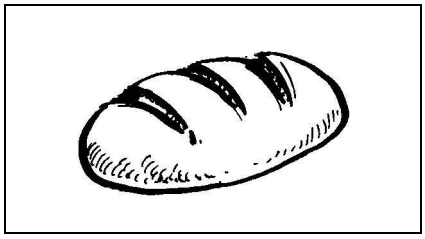


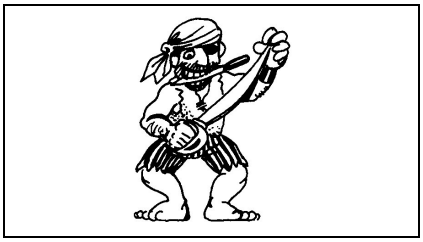






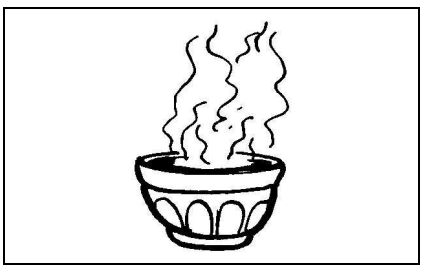




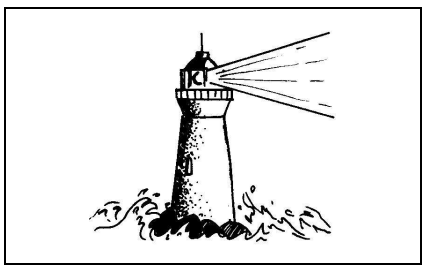


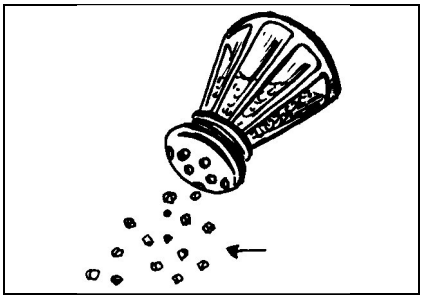


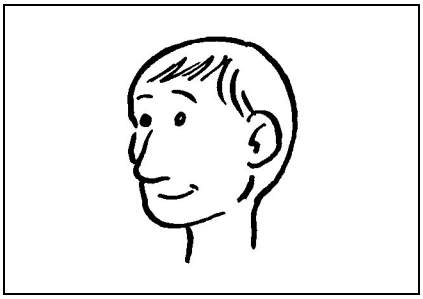

Mr Pahl
7 rue du Parc
15075 GIROUX

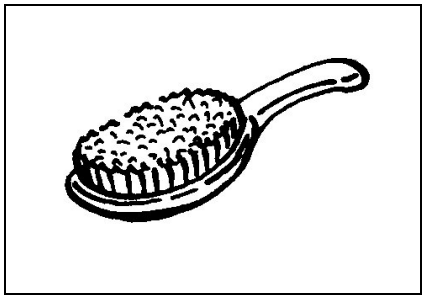


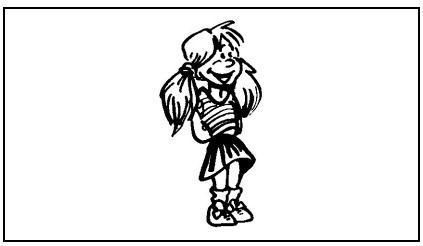


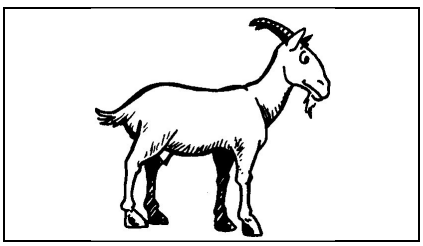


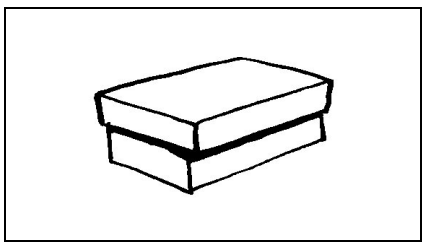












Tâche de détection de l'intrus phonémique

Passation : collective

Consigne :

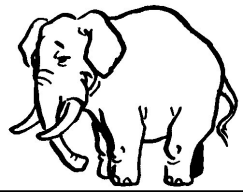
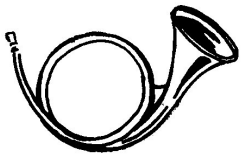
Regardez ces images. Je vais vous donner chaque fois trois mots et vous allez entourer l'image du mot qui ne finit pas par le même son que les deux autres. Je vais vous expliquer encore plus par des exemples.

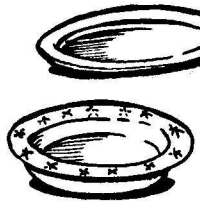
شوفو التصاور هذوما ، آنا في كل مرّة بش نقلكم ثلاثة كلمات و انتوما بش ادوّرو التصويرة إلى اسمها ما يوفاش بنفس الصوت هي و لخرين – وهاو بش نجربو مع بعضنا

ساق - بوق - فيل

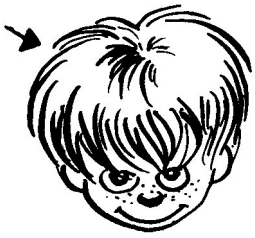
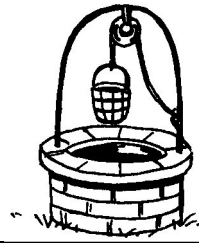
Matériel :

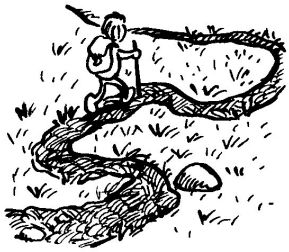
غصن	**ذئب**	**صحن**
خبز	**بئر**	**شعر**
كتاب	**غروب**	**طريق**
برميل	**عنقود**	**غربال**
حليب	**شروق**	**لصاق**
رمح	**سطح**	**صفر**

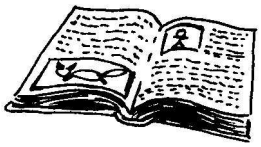



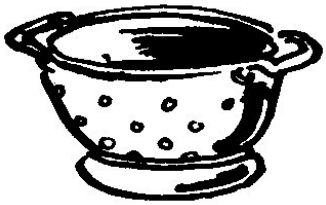
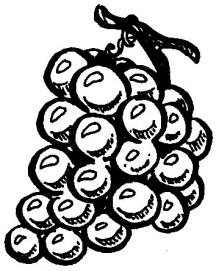
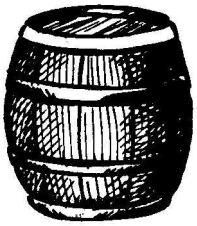



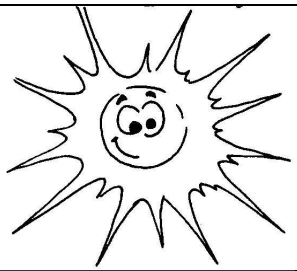


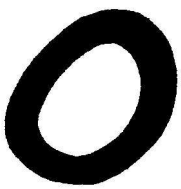
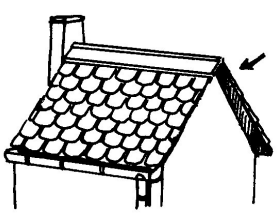
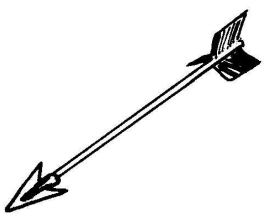





Epreuves d'évaluation des connaissances morphologiques

Tâche de jugement des liens morphologiques

Passation : collective

Consigne :

Je vais vous donner chaque fois une série deux mots. Il y a des mots qui sont de la même famille et d'autres qui ne sont pas de la même famille. Lorsque les mots sont de la même famille vous allez entourer l'image devant vous. Lorsque vous trouvez qu'ils ne sont pas de la même famille, vous mettez une croix sur l'image. Je vais vous expliquer encore plus par des exemples.

بش نقرالكم كل مرَة زوز كلمات. الكلمات هاذوما ينجمو يكونو من نفس العايلة وإلاّ لا. كان تشوفو إلّي هوما من نفس العايلة تدوَرو التصويرة إلّي قدَامكم. كان تشوفو إلّي هوما مش من نفس العايلة تعملو قاطع و مقطوع على اتصويرة. و هاو بش نجرَبو مع بعضنا

Matériel :

1- كرسي ـ كرَاسة
2- حيوان ـ حيوانات
3- قرا ـ قراية
4- مثلّث ـ مربَع
5- بحر ـ حبار
6- اسم ـ لقب
7-طفل ـ طفلة
8 -سطَر ـ مسطرة
9 -يسكن ـ سكَينة
10- ليل ـ رقد
11-بكى ـ يبكي
12-طيَارة ـ مطار
13-شركة ـ شكارة
14-سبَورة ـ طباشير
15-كلا ـ ياكلو
16-خبزة ـ خبَاز

Tâche morpho-syntaxique

Passation : collective

Consigne :

Je vais vous lire chaque fois une phrase. Cette phras peut être correcte ou incorrecte. Si vous trouvez qu'elle est correcte, vous entourez l'image devant vous (montrer). Si vous trouvez qu'elle est incorrecte, vous devez faire une croix sur l'image. Je vais vous expliquer encore plus par des exemples.

بش نقرالكم في كل مرّة جملة، الجملة هاذي تنجّم تكون صحيحة و إلاّ غالطة. كان تشوفو الي الجملة صحيحة تدورو التصويرة إلّي قدّامكم. كان تشوفو الي الجملة غالطة تعملو قاطع و مقطوع على التصويرة. و هاو بش نجرَبو مع بعضنا

Matériel :

Jugement de l'ordre

سامي شرب كاس حليب
الطفلة الصغيرة تمشي أمها مع
القطّوس وراء يجري الفار
الوليّد قرا كتاب
المعلّم خرج من القسم
الأرنب تجري الغابة في
المرا خيّطت روبة
محمّد شرى شكلاطة
مريم الشبّاك حلّت
أمّي خبّات الفلوس المخدّة تحت

Jugement de la morphologie des mots

أحمد كلات تفّاحة
سامي شرب كاس حليب
الوليّد روّح من المكتب
التلامذة لعب في السّاحة
جمع صالح الزيتون
محمد واقفين بحثا الدّار
العصافر هرب من االقفص
البنيّة قعدت فوق الكرسي
الكرطابة متاعي جديدة
سلمى لبس حوايج العيد

Epreuve de connaissance des lettres

Passation : collective

Consigne :

Maintenant, on va lire des lettres. Vous devez entourer chaque fois la lettre que je vais nommer dans l'encadré. Je vais vous expliquer encore plus par des exemples. Mettez le doigt sur l'image du bateau. Maintenant, entourez la lettre ب.

توّا بش نقراو حروف. لازمكم في كل مرّة تدوَرو الحرف إلّي نقلكم عليه في المربّع. و هاو بش نجرَبو مع بعضنا

حطّو صوابعكم على تصويرة الفلوكة. توّا دوّرو حرف ب

حطّو صوابعكم على تصويرة الكرهبة. توّا دوّرو حرف س

توّا بش نبداو

حطّو صوابعكم على تصويرة الوجه إلّي يضحك. توّا دوّرو حرف **ر**

حطّو صوابعكم على علامة مَعَ. توّا دوّرو حرف **ع**

حطّو صوابعكم على تصويرة الطيارة. توّا دوّرو حرف **ج**

حطّو صوابعكم على تصويرة القلب. توّا دوّرو حرف **ق**

حطّو صوابعكم على تصويرة الأمبوبة. توّا دوّرو حرف **د**

حطّو صوابعكم على تصويرة النوارة. توّا دوّرو حرف **ط**

حطّو صوابعكم على تصويرة الشجرة. توّا دوّرو حرف **ث**

حطّو صوابعكم على تصويرة الكرسي. توّا دوّرو حرف **ص**

حطّو صوابعكم على تصويرة الجواب. توّا دوّرو حرف **أ**

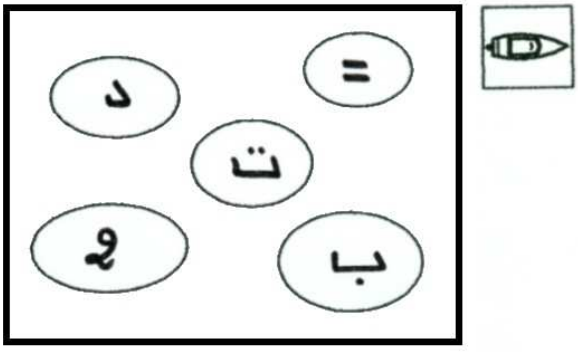
حطّو صوابعكم على تصويرة البسكلات. توّا دوّرو حرف **ه**

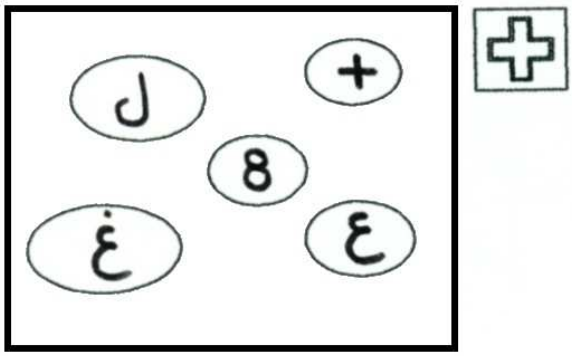
حطّو صوابعكم على تصويرة النجمة. توّا دوّرو حرف **ك**

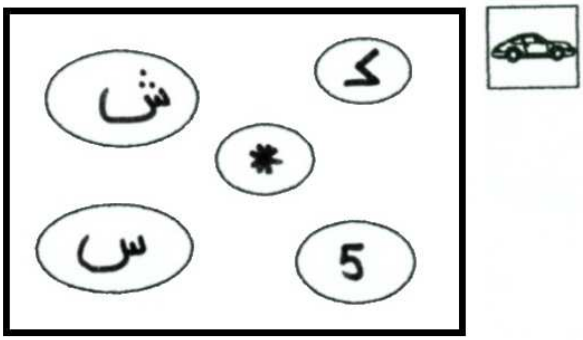
حطّو صوابعكم على تصويرة السهم. توّا دوّرو حرف **ن**

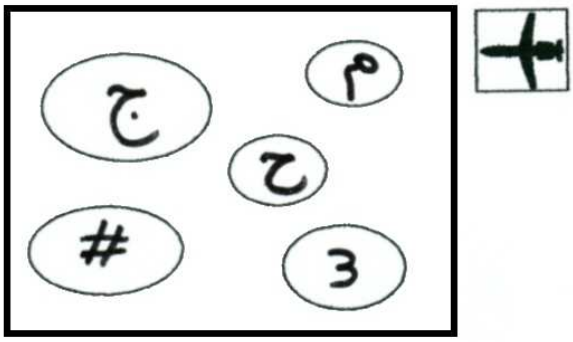
Matériel :

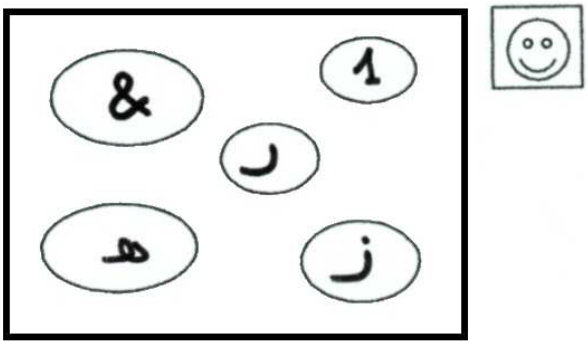
اليوم :
الإسم واللقب :
تاريخ الولادة :

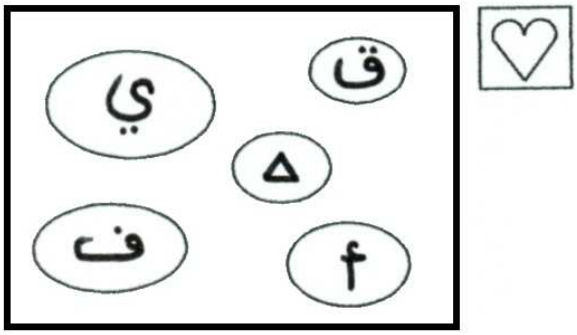


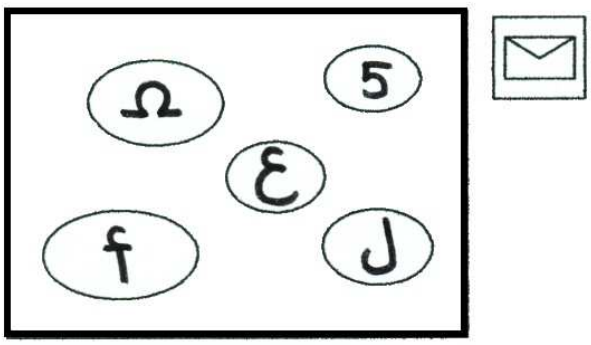


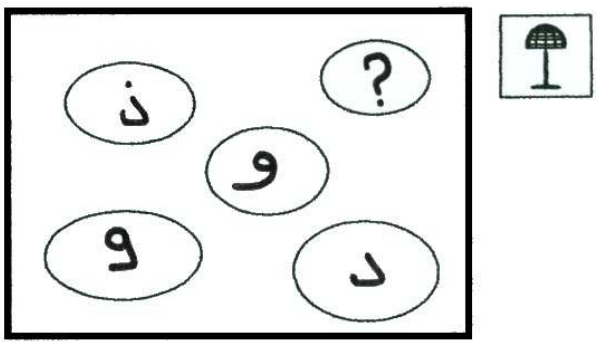


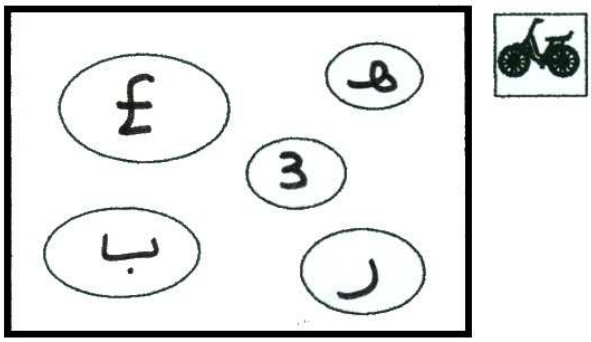


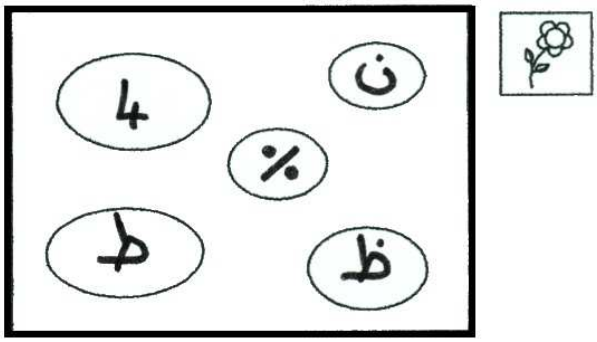


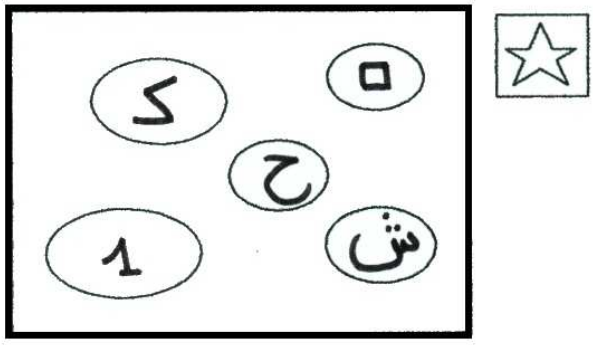


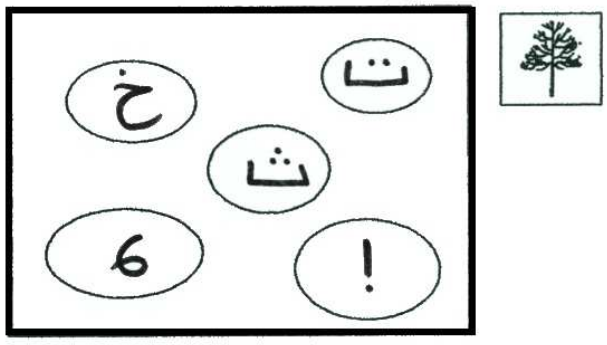

5
Ω
ع
f
ل

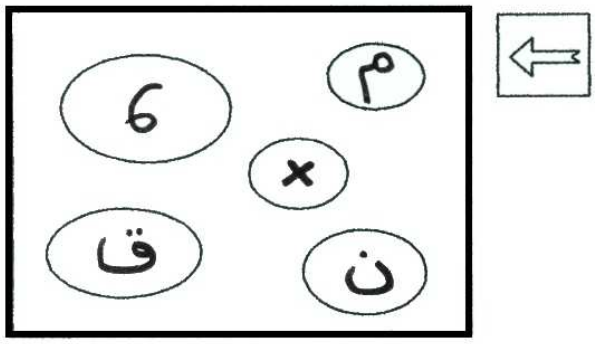

?
ن
و
و
د

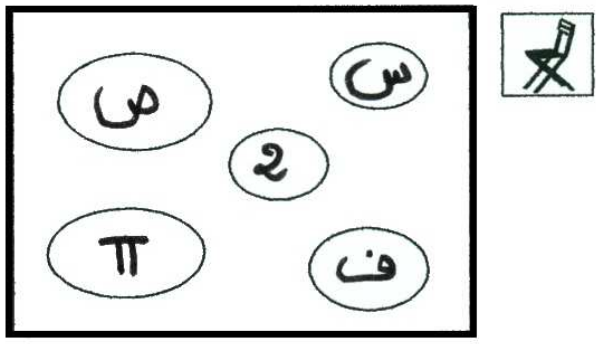

£
ه
3
ب
ر


4
ن
%
ط
ظ


ك
ם
ح
1
ش


خ
ت
ث
6
!


6
م
x
ق
ن


ص
س
2
π
ف

Annexe II : Questionnaire d'évaluation des compétences et comportements scolaires destiné aux enseignants

Ce questionnaire a été conçu par Agnès FLORIN, Philippe GUIMARD & Isabelle NOCUS[97]. Il vise à aider les enseignants à repérer les modalités habituelles de fonctionnement des enfants à l'école maternelle, et à leur donner un outil de dialogue avec leurs partenaires de l'école. Nous avons traduit ce questionnaire en langue arabe et nous l'avons administré aux enseignants de la classe préparatoire (mai-juin 2005).

استـــجـــواب خـــاص بـتـلامـيـذ الـسـنـة الـتـحـضـيـريـة

الإســم و اللــقب	:	**الــجنس**	: .....
تــاريخ الــولادة	:	**العــمر**	: .....
تــاريخ الإجــابة	:	**القــسم**	: .....

الرّجاء الإجابة على كلّ الأسئلة دون إستثناء ، أبرز في كلّ سؤال السلوك الحالي للطفل أو السلوك الأكثر حدوثا.

1- بالنسبة إلى المحادثة في القسم، هل يمكن القول أنّه(ها) يشارك :

- كـثـيـــرا. ..... ☐
- بـصـفـة مـتوسّطـة. ..... ☐
- قـلـيلا أو لا يـشـارك أبـدا. ..... ☐

2- بالنسبة إلى المحادثة في القسم، هل يمكن القول أنّه يشارك:

- أحيانا برويّـة ..... ☐
- نادرا برويّة (لا يندمج كثيرا في سياق الحديث). ..... ☐
- من المستحيل التكهّن بذلك (غير مفهوم أو لا يتكلّم وسط المجموعة) ..... ☐

3- بالنسبة إلى اللغة (مجموع الكلمات)، و مقارنة بالمستوى العادي لقسمك، هل ترى أنّه :

- حـسـن ..... ☐
- مـتـوسط ..... ☐
- غـير كاف ..... ☐

4- بالنسبة إلى اللغة (النحو و الصرف) ، و مقارنة بالمستوى العادي لقسمك ، هل ترى أنّه :

- حـسـن ..... ☐
- مـتـوسط ..... ☐
- غـير كاف ..... ☐

5- بالنسبة إلى اللغة (النطق) ، و مقارنة بالمستوى العادي لقسمك ، هل ترى أنه :

[97] En accès libre sur: www.lettres.univ-nantes.fr/labecd/index.htm.

- حـســـن..... ☐
- مــتــوسط..... ☐
- غــير كاف..... ☐

6- إذا ناديته بإسمه ، و أنت تدير ظهرك له ، قصد إعطاءه بعض التعليمات البسيطة أو لتسأله عما هو بصدد فعله ، هل يمكن القول أنه :

- يظهر أنه فهم المقصود من كلامك ..... ☐
- تشكّ أنه فهم ذلك ..... ☐
- لا يظهر أدنى إجابة ، سواء كانت شفاهية أو لا..... ☐

7- هل بإستطاعته الإنتباه ،

- بصفة منتظمة و دائمة ..... ☐
- بصفة غير منتظمة، قليلة الدّوام..... ☐
- يعجز غالبا عن تركيز انتباهه ..... ☐

8- عند قيامه بنشاط ما ، هل يظهر بعض الإستقلالية :

- بتصرّف وحدّة ..... ☐
- يبحث عن المساعدة ..... ☐
- لا يبحث عن المساعدة لكن سرعان ما تقع مجاوزته ..... ☐

9- عند قيامه بعمل ما ، هل هو :

- سريع ..... ☐
- ليس سريعا و لا بطيئا ..... ☐
- بطيء ..... ☐

10- بالنسبة إلى الأنشطة التي تعتمد على الذاكرة (أناشيد ، قصص ....) هل يمكن القول أنّ له : ☐

- ذاكرة جيّدة .....
- بعض الصعوبات ..... ☐
- صعوبات خطيرة ..... ☐

11- خلال الأنشطة المدرسيّة ، هل تظهر عليه علامات التعب :

- قــلــيــلا ..... ☐
- من وقت لآخر أو بالنسبة إلى أنشطة معيّنة ..... ☐
- أحيانا ..... ☐

12- حسب رأيك، حتّى يتابع النسق في القسم ، هل يمكن القول أنه :

- لا يجد أيّة صعوبة (يتابع النسق بصفة عادية) ..... ☐
- يحتاج للمساندة و التشجيع ..... ☐
- بحاجة للإجبار و تذكير بذلك ..... ☐

13- عند قيامه بعمل ما ، هل يمكن القول أنّه :

- يعرف كيف ينظّم عمله (يتوقع سابقا المراحل التي يجب توخيّها) ..... ☐
- ينسّق عمله تدريجيّا حسب التعليمات ..... ☐
- لا يستطيع تنظيم عمله (لا ينفّذ إلاّ إذا أُمر) ..... ☐

14- في الأنشطة المدرسية ، هل يظهر :

- ثقة في النفس مبرّرة (يحسّ بأنه قادر على القيام بذلك و يحسن التصرف) ..... ☐

- ثقة عمياء في النفس (لا يحسن تقدير الصعوبات و امكانياته) ..... ☐
- عدم الثقة في النفس (من المتوقع أن يتنازل) ..... ☐

15- في ما يخصّ الحياة في القسم ، هل يمكن القول أنه :

- يحترم بدون صعوبة قواعد الحياة الجماعية ..... ☐
- يجد بعض الصعوبات في احترام القواعد الثابتة ..... ☐
- لا يحترم قواعد الحياة إلاّ إذا كان مجبرا على ذلك ..... ☐

16- بالنسبة إلى الأنشطة التي تعتمد على اتقان الحركات (الأشكال ، الأعمال اليدوية) :

- يتقنها جيّدا ..... ☐
- يجد بعض الصعوبات بالنسبة إلى أنشطة معيّنة ..... ☐
- يجد صعوبات خطيرة ..... ☐

17- هل يشارك في الأنشطة الحركية – النفسية :

- بإتقان ..... ☐
- يجد بعض الصعوبات بالنسبة إلى أنشطة معينة ..... ☐
- يجد صعوبات خطيرة ..... ☐

18- بالنسبة إلى الأنشطة التي تعتمد على التفكير المنطقي (حلّ المسائل ، أنشطة منطقية) :

- يحسن استعمال استراتيجيات التلمّس دون مساعدة خارجية ..... ☐
- يحسن استعمال استراتيجيات التلمس بمساعدة خارجية ..... ☐
- يجد صعوبة في استعمال الإستراتيجيات ..... ☐

19- بالنسبة إلى الأنشطة التي تعتمد على مقاربة الأعداد (مقارنة ، فرز ، تنظيم الأشياء) هل يمكن القول أنّه :

- يتقنها جيّدا ..... ☐
- يجد بعض الصعوبات بالنسبة إلى أنشطة معينة : ..... ☐
- يجد صعوبات خطيرة ..... ☐

20- إذا أخذنا بعين الإعتبار مستواه الحالي ، هل ترى أنه في السنوات الدراسية المقبلة :

- لن يجد أيّة صعوبة في تعلّم القراءة ..... ☐
- سيجد بعض الصعوبات الطفيفة ..... ☐
- سيجد صعوبات خطيرة ..... ☐

Corrélation entre le score global au questionnaire et les épreuves langagières et cognitives en CP

	Score global au questionnaire
Comptage syllabique	**.45***
Segmentation syllabique	**.43***
Suppression syllabique	**.47***
Intrus syllabique	.32*
Intrus rime	.14
Morpho-syntaxe	**.45***
Connaissance des lettres	**.52***
Compréhension orale	.27
Compréhension consignes	**.43***
Dénomination rapide	-.26
Intelligence	.37*
MCT	**.42***
Graphisme	**.55***
Concepts de base	**.49***
Attention	.39

*s à p<.01
En gras : les corrélations les plus fortes.

Analyse de régression entre le score global au questionnaire et aux mesures langagières à T1 et les performances à l'épreuve IME2

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
SG questionnaire	.310	.212	.227	48.392
Connaissance des lettres	.566	.265	.276	32.066
Segmentation syllabique	.176	.287	.304	24.059
Comptage syllabique	.388	.309	.321	19.406

Analyse de régression entre le score global au questionnaire et aux mesures langagières à T1 et les performances à l'épreuve Timé2

Variables explicatives	**β**	**R^2**	**R^2 ajusté**	**F**
SG questionnaire	.353	.272	.287	68.824
Connaissance des lettres	.302	.365	.386	52.354
Détection intrus syllabique	.159	.407	.424	37.920

Estimation de la variance expliquée des items du questionnaire

R^2	Estimation
Q1 Quantité de participation en classe	.671
Q2 Qualité de participation en classe	.670
Q3 Vocabulaire	.856
Q4 Syntaxe	.726
Q5 Prononciation	.645
Q6 Compréhension	.641
Q7 Attention	.643
Q8 Autonomie	.636
Q9 Rapidité	.584
Q10 Mémoire	.686
Q11 Fatigue	.347
Q12 suivi du rythme	.841
Q13 Organisation	.705
Q14 Confiance en soi	.634
Q15 Respect des règles	.443
Q16 Maîtrise de gestes	.744
Q17 Activités EPS	.745
Q18 Raisonnement	.737
Q19 Approche du nombre	.765

Annexe III : Les épreuves de lecture en 2ème année primaire (T3)

En mai 2007, les enfants ont passé trois épreuves de lecture : une épreuve d'identification de mots écrits en contexte imagé (IME) adapté de Khomsi (2004) et un test d'identification de mots écrits inspirée du Timé2 (Ecalle, 2003). Pour l'évaluation de la compréhension en lecture, nous avons utilisé une adaptation de l'épreuve C.20 (Khomsi, 1990). Avant chaque pssation, les enfants sont invités à remplir l'encadré en haut à droite qui comporte des informations sur la date de passation, le nom, l'âge et la classe de l'élève.

Epreuve d'identification de mots écrits (IME) :

Passation : collective

Consigne :

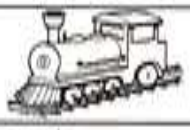
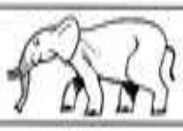
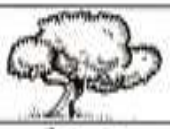
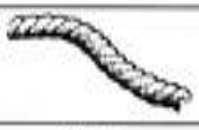
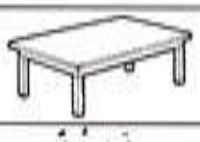
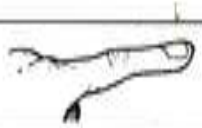
Vous avez devant vous des feuilles sur lesquelles il y a des dessins. Au dessous de chaque dessin, il y a une étiquette avec un mot écrit. Certains mots vont bien avec les dessins et d'autres pas. Certains mots sont bien écrits et d'autres ont des fautes. Quand les mots ont des fautes ou quand ils ne vont pas bien avec les dessins, vous barrez le tout avec deux grands grands traits.

شفتو التصاور إلي قدَامكم. تحت كل تصويرة فمَ كلمة. الكلمة هاذي تنجَم تمشي مع التصويرة و الاَ لا. واتنجَم تكون صحيحة و إلاَ لا معناها غالطة. انت كل ما تلقى كلمة موش صحيحة و الاَ ما تمشيش مع التصويرة تحط علامة قاطع و مقطوع. هاو امثلة بش تفهم أكثر

Matériel :

Nous ne présentons que la première partie de cette épreuve.

اليوم :	الإسم واللّقب :	القسم :

 زَهْرَةٌ	 مِفْظَةٌ	 سَيَّارَةٌ	 سَمَكَةٌ
 سَطْلٌ	 سِرْوَالٌ	 دُبٌّ	 كَلْبٌ	 طِفْلٌ
 سَحَابٌ	 نَـارٌ	 بُومَةٌ	 بَـرَةٌ	 شَجَرَةٌ
 وَرْدَةٌ	 أَرْبَعَةٌ	 شَمْسٌ	 أَسَدٌ	 كَأْسٌ
 فَأْرٌ	 نَمْلَةٌ	 عَيْنَـانِ	 حَبْلٌ	 فِنْجَانٌ
 نَجْمَةٌ	 رِجْلٌ	 طَاوِلَةٌ	 تُفَّاحَةٌ	 دَجَـاجَةٌ
 يَدٌ	 سُلَّمٌ	 عُصْفُورٌ	 سَاقٌ	 بِنْتٌ

Test d'identification de mots écrits (Timé2) :

Passation : collective

Consigne :

Vous avez devant vous une feuille composée de plusieurs images et mots. Chaque fois, vous devez trouvez le mot écrit correctement. Je vais vous expliquer encore plus par des exemples.

شفتو الورقة إلّي قدّامكم، فيها تصاور و كلمات. في كل مرّة لازمكم تلقاو الكلمة إلّي مكتوبة بالصحيح هاو امثلة بش تفهم أكثر

Tâche 1 :

Au début, vous devez entourer, parmi les cinq mots, celui que vais énoncer. Je vais vous expliquer encore plus par un exemple.

مالأول لازمكم تدوّرو ما بين الخمسة كلمات، الكلمة إلي آنا بش نقولهالكم– وهاو بش نجربو مع بعضنا

Tâche 2 :

Maintenant, vous travaillerez seuls. Vous devez entourer, parmi les cinq mots, celui qui est écrit correctement et qui correspond à l'image. Je vais vous expliquer encore plus par un exemple.

توّا بش تخدمو وحدكم. لازمكم تدوّرو ما بين الخمسة كلمات الكلمة إلّي مكتوبة بالصحيح و لّي تمشي مع التصويرة– وهاو بش نجربو مع بعضنا

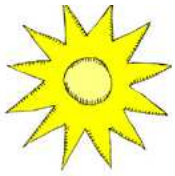
Tâche 3 :

Maintenant, vous allez entourer le mot écrit correctement et qui va bien avec le mot écrit dans l'encadré (montrer avec les doigts). Je vais vous expliquer encore plus par un exemple.

توّا بش تدورو الكلمة إلّي مكتوبة بالصحيح ولّي تمشي مع الكلمة المكتوبة في المستطيل الفوق– وهاو بش نجربو مع بعضنا

Matériel :

الإسم و اللّقب: تاريخ الولادة: القسم: اليوم:
النّشاط الأوّل: النّشاط الثاني: النّشاط الثّالث: المجموع:



شَمَسٌ
شَمْصٌ
شَمْسٌ
شَمْعٌ



تُرَابُنْ
تَرُابٌ
تُرَاتٌ
تُرَاثٌ

أَرْنَبٌ

جْرْزٌ
جَزَرٌ
جَزَرُنْ
جَرَسٌ



صْتُوَتٌ
صَوْمٌ
صَوْتٌ
صَوُتٌ



شَارِغٌ
شَاعِرٌ
شَارِعُنْ
شَارِعٌ



سَطْلٌ
سَطِلٌ
سَطْحٌ
سَتْلٌ



لَوْزُنْ
لَوَزٌ
لْزَوٌ
لَوْنٌ



حِسَاتٌ
حْبُاسٌ
حِسَابٌ
حِصَابٌ



قِرْطٌ
قْدُرٌ
قِرْدُنْ
قِرُدٌ



غْبَارٌ
غُارْبٌ
غُرَابٌ
غُزَابٌ



فَرُسٌ
فَرَسٌ
فَرَصٌ
فَرَحٌ



سَاحِرٌ
سَاحِزٌ
سَاحِرُنْ
سْحَارٌ



رِيحُنْ
رِيخٌ
رْيَحٌ
رِيحٌ



لِصَاقٌ
لِسَانٌ
لْسُانٌ
لِشَانٌ



قْرِمٌ
قَمُرٌ
قَمَرُنْ
قَصْرٌ

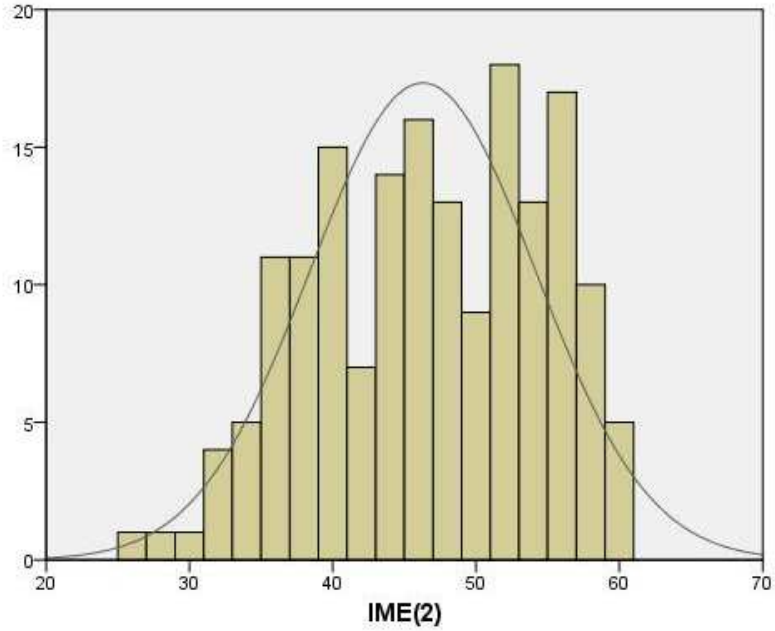
1 وَاحِدٌ وَاحِدٌ وَاحِدَانْ وُحَادٌ وَاحَةٌ	عِلْمٌ عَلَمَانْ عَلَمٌ عَمَلٌ عِلْمٌ	جِرْسٌ جِسْرٌ جِسْمٌ جِشْرٌ جِصْرٌ
مَوْجٌ مَوْزَانْ مَزْوٌ مَوْزٌ	نَهْرٌ نَهْجٌ نَهْرٌ نَهْرَانْ	بَيْضٌ بَيْذٌ بَيْتٌ بَيْضٌ
فَأْسٌ فَأْرٌ فَأْسٌ فَأْسٌ فَأْسٌ	بَصَلٌ بَسَلٌ بَصَلٌ بَلَصٌ بَطَلٌ	سَهْلٌ سَهْمٌ سَهَمٌ سَهْمَانْ سَمْهَمٌ
ثَوْرَانْ ثَوْرٌ ثَوْرٌ ثَرْوٌ ثَوْبٌ	جَبَلٌ جَلَبٌ جَمَلٌ جَبَلٌ جَبَلَانْ	شَارِبَانْ شَارِبٌ شَرَابٌ شَارِبٌ شَرَابٌ

تِلْمِيذٌ قَسَمٌ قَشَمٌ قَسَطٌ قَمَسٌ قَصَمٌ	تَمْرٌ نَخْلَةٌ نَخْلَةٌ نَخْلَةٌ نَتَخَلٌ نَخْلَتَانْ	كُرَةٌ مَرْسَى مَرْمَى مَرْمَا مَرْمَى مَمْرَى	مِحْرَاثٌ حَلَقٌ حَقَلٌ حَقَلٌ حَقَلَانْ حَقَلٌ
كِتَابٌ صَفْخَةٌ صَحْفَةٌ صَفْحَةٌ صَحْفَةٌ صَفْحَتَانْ	وَرْدَةٌ رَبِيعَانْ رَبِيعٌ رَضِيعٌ رَبِيعٌ رَبِيعٌ	حَيَوَانٌ قَنَطَةٌ قَصَبَةٌ قَطَةٌ قَنَةٌ قَطَةٌ	لَوْنٌ أَسَدٌ أَسَدٌ أَسْوَدٌ أَصْوَدٌ أَسْدٌ
قَفَّةٌ سُوقٌ سُوَقٌ سُورٌ سُوقَانْ سُقُوٌ	بَحْرٌ رَمْلَانْ رَمْلٌ رَمْلٌ رَجْلٌ رَمْلٌ	مُشْطٌ شَعَرٌ شَرْعٌ شَعْرٌ شَعْرَانْ شَعْبٌ	غِنَاءٌ مَسْبَحٌ مَصْرَحٌ مَرْحَلٌ مَسْرَحٌ مَسْرَحٌ

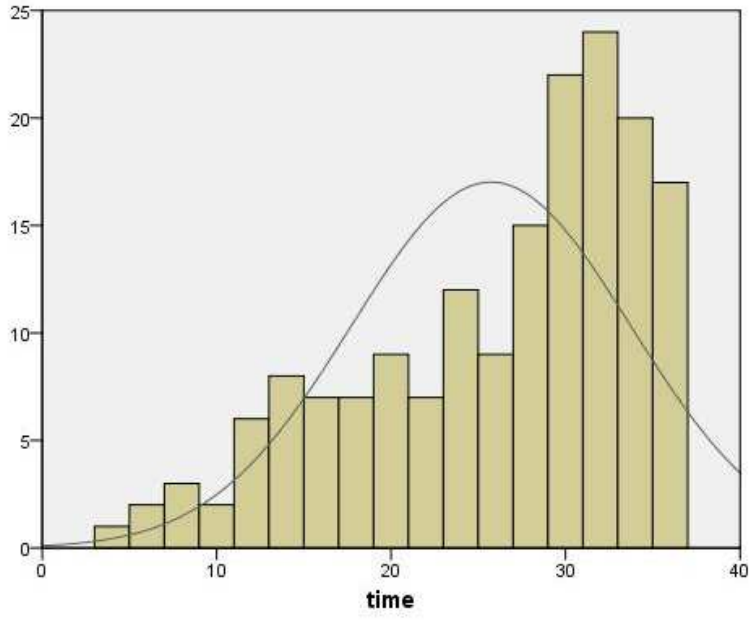
Centile 25 et centile 75 des épreuves de reconnaissance de mots en 2ème année (IME2 et Timé2)

	Centile 25	**Centile 75**
IME2	40	53
Timé2	20	32

Distribution des scores à l'épreuve IME2



Distribution des scores à l'épreuve Timé2



Epreuve de compréhension écrite :

Passation : collective

Consigne :

Vous avez devant vous des séries de quatre images. Chaque fois, vous allez lire la phrase écrite au dessous et entourez, parmi les quatre images, celle qui va avec.

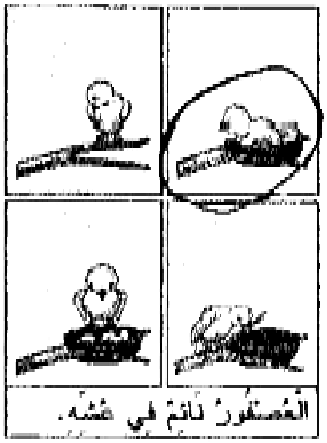
قدامكم مجموعات متاع تصاور. كل مجموعة فيها أربعة تصاور. في كل مرة إنتوما بش تقراو الجملة المكتوبة اللوطة وتدورو التصويرة إلي تمشي معاها. وهاو بش نجربو مع بعضنا

Matériel :

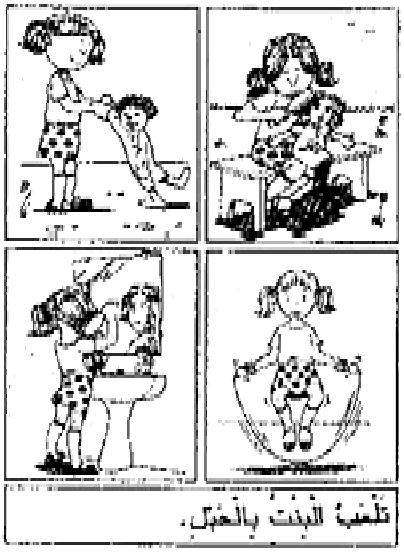
Nous ne présentons que la première partie de cette épreuve.

الإسْم وَ اللَّقَب :

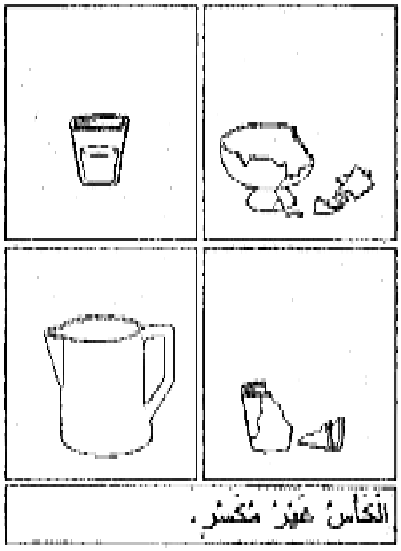
القِسْم :



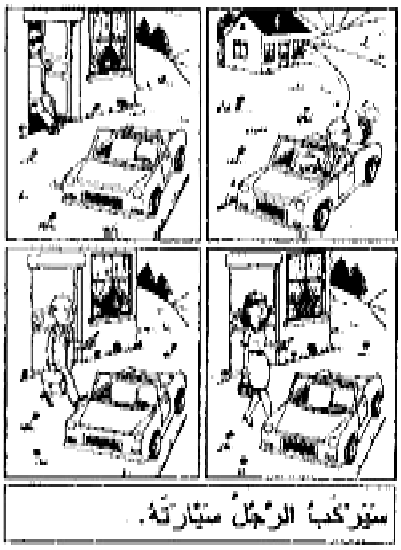
العُصفورُ نائمٌ في عُشّه.



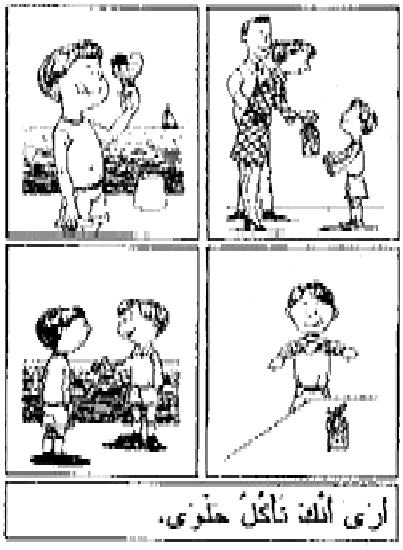
تلْعبُ البنتُ بالحَبْل.



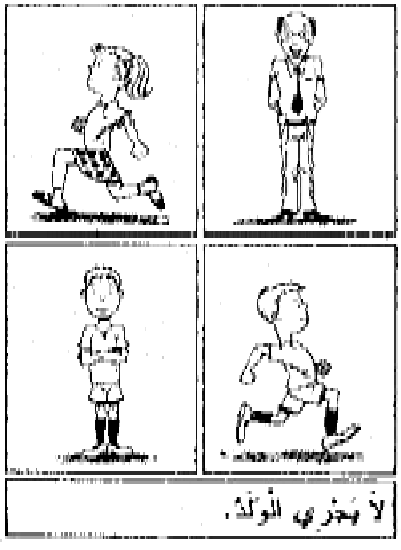
الكأسُ غيرُ مُكسّر.



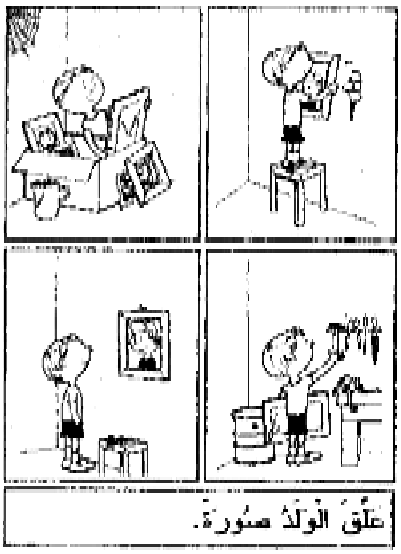
سيَرْكَبُ الرّجُلُ سيّارَتَه.



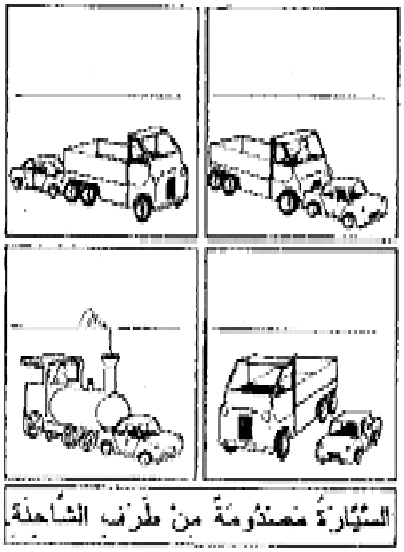
أرى أنّكَ تأكُلُ حلْوى.



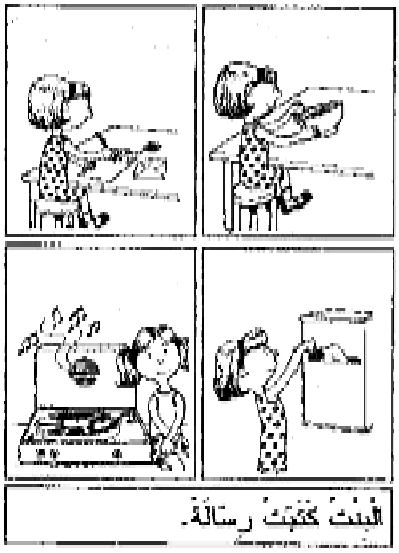
لا يجْري الوَلَدُ.



عَلّقَ الوَلَدُ صُورَةً.



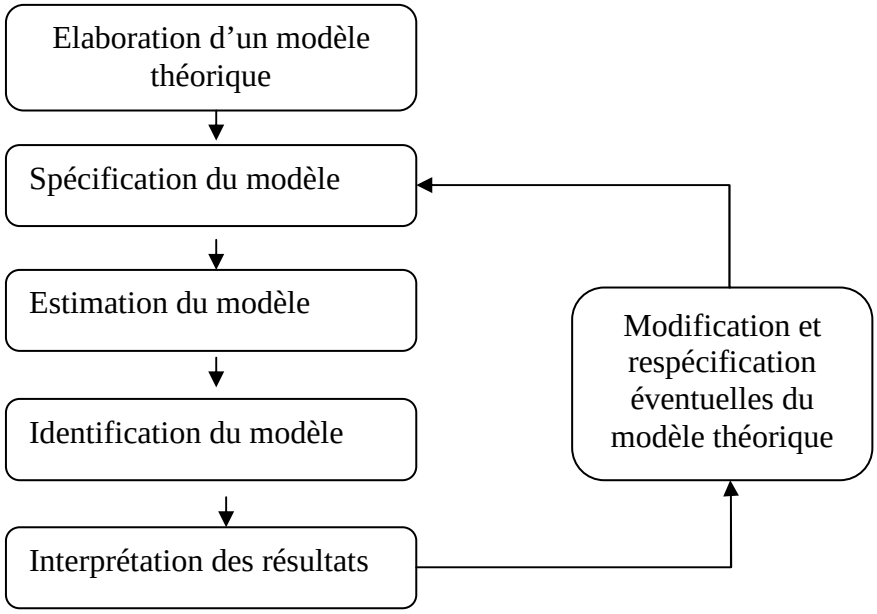
السّيّارَةُ مَصْنُوعَةٌ مِنْ مِقْرَفِ الشّاحِنَة.



البِنْتُ كَتَبَتْ رِسالَةً.

Annexe IV : Démarche méthodologique pour les modèles d'équations structurelles

Un modèle d'équations structurelles teste les relations linéaires hypothétiques entre plusieurs variables. Il s'applique aux études confirmatoires ou recherches hypothético-déductives.
Pour tester un modèle structurel, cinq étapes sont nécessaires : (1) la construction d'un modèle thérique, (2) la spécification du modèle, (3) l'estimation du modèle, (4) l'identification du modèle, (5) le test du modèle et éventuellement la modification du modèle. Nous présentons une synthèse de la description des différentes étapes fournie par Roussel & al. (2002).



Démarche méthodologique de développement et de test d'un modèle d'équations structurelles (d'après Roussel & Wacheux, 2006)

1. Développer un modèle théorique

Un modèle théorique correspond à un ensemble de relations proposant une explication cohérente et compréhensible d'un phénomène. Chaque relation entre variables est soutenue par une hyopthèse qui s'appuie sur un cadre théorique et/ou des observations empiriques. Celles-ci se réfèrent aux observations relevées soit dans des recherches ou des études antérieures (démarche déductive), soit au cours d'une phase initiale dite exploratoire pendant laquelle des observations permettent d'élaborer des hypothèses sur les relations entre

variables (démarche inductive).

2. Spécification du modèle

La spécification d'un modèle théorique consiste d'une part à déterminer le modèle de mesure des variables latentes et d'autre part à déterminer le modèle structurel. Au total, le modèle complet se divise en deux sous-modèles : le modèle de mesure et le modèle structurel. Le *modèle de mesure* décrit l'ensemble des relations impliquées par les variables mesurées. Le *modèle structural* décrit les relations impliquant exclusivement les variables latentes.

Au cours du processus de spécification, le concepteur du modèle doit déterminer quels seront les paramètres fixes, libres et contraints. Les paramètres sont les coefficients de corrélations ou de covariance, de régression, mais aussi les contributions factorielles des indicateurs aux variables latentes. Les paramètres fixes désignent les liens potentiels qui ne seront pas estimés dans le modèle. Ils seront fixés à 0 pour ne pas entrer dans les procédures de calcul. Les paramètres libres désignent les paramètres dont la valeur est inconnue et que l'on cherche à estimer. Enfin, les paramètres contraints désignent les paramètres libres auxquels le modèle impose une contrainte soit d'égalité entre deux variables, soit de supériorité ou d'infériorité de la valeur d'une variable par rapport à la valeur d'une autre ou par rapport à un seuil.

3. Estimation du modèle

L'objectif est de proposer des estimations de chaque paramètre du modèle de façon à ce que les valeurs de ces paramètres produisent une matrice de covariance plus semblable possible à la matrice de covariance générée à partir des données observées.

Il existe différentes méthodes d'estimation. La plus robuste et la plus répandue est celle du maximum de vraisemblance (Maximum Likelihood).

4. Identification du modèle

L'identification d'un modèle théorique signifie qu'il est possible de trouver une solution unique. L'identification du modèle est l'estimation par le logiciel des valeurs des paramètres dans les équations structurelles. L'identification dépend du nombre de degrés de liberté dans le modèle. Les degrés de liberté représentent l'écart entre entre le nombre de corrélations ou de covariances estimés et potentielles.

Le nombre de degrés de liberté dans le modèle est le nombre de paramètres non estimés (parmi les paramètres potentiellement estimables, compte tenu du modèle et des données).

Pour éviter les problèmes de sous-identification, des conditions sont utiles à savoir :

- Prévoir, si possible, un minimum de 3 indicateurs par variable latente.
- Fixer à un la variance des variables latentes. Il est souhaitable pour la mesure de chaque variable latente de fixer le premier paramaètre à 1 (la contribution de son premier indicateur=1).
- Tester un modèle de départ, le plus parcimonieux possible, puis complexifier au besoin. Cette approche consiste à estimer des modèles alternatifs en se limitant aux relations uni-directionnelles dans un premier temps puis en introduisant des effets réciproques dans un second temps.

5. Interprétation des résultats

L'évaluation d'un modèle d'équations structurelles consiste à apprécier la qualité de l'ajustement du modèle aux données empiriques. Il existe de nombreux indices d'adéquation.

Tableau récapitulatif des valeurs généralement admises pour les principaux indices d'ajustement

Indice	**Valeur**
Indices absolus	
$\chi 2$	Non significatif
GFI et AGFI	>0.9 (valeurs comprises entre 0.7 et 0.8 sont tolérées pour le cas des modèles complexes)
RMR	La plus proche de 0
RMSEA	<0.08 et si possible <0.05
Indices incrémentaux	
NFI	>0.9
IFI	>0.9
CFI	>0.9
Indices de parcimonie	
$\chi 2$ ajusté aux degrés de liberté	La plus faible (ne pas dépasser le 5)
AIC, ECVI	La plus faible possible (comparaison)
PGFI	La plus faible possible (comparaison)

6. La respécification éventuelle du modèle

La modification du modèle théorique intervient lorsque l'on souhaite améliorer sa qualité d'ajustement aux données empiriques ou son adéquation aux hypothèses qui le sous-tendent. Elle consiste à ajouter ou supprimer des paramètres estimés du modèle en se basant sur des arguments théoriques. Les logiciels proposent des indicateurs facilitant l'identification des variables concernées par la respécification. Le premier indicateur est celui des *résidus standardisés ou normalisés* de la matrice des covariances ou des corrélations. La valeur du résidu se lit alors comme un score z, on la considère comme anormalement élevée si elle est supérieure à 1,96 ou 2,58. Il est possible d'examiner un deuxième type d'indicateur : les indices de modification aidant à la respécification du modèle.

Etude longitudinale prédictive de la réussite ou de l'échec dans l'apprentissage de la lecture en langue arabe

RÉSUMÉ

Cette étude longitudinale a pour objectif principal d'étudier la valeur prédictive globale et individuelle des compétences langagières et cognitives précoces sur les performances en lecture en langue arabe. Il s'agit d'autre part de déterminer la capacité d'un questionnaire d'évaluation des comportements et compétences scolaires renseigné par les enseignants en classe préparatoire (CP) à prédire le niveau ultérieur en lecture. Les résultats obtenus auprès de 171 enfants tunisiens suivis de la classe préparatoire à la 2ème année primaire confirment la valeur prédictive de la connaissance des lettres et des connaissances phonologiques. Elles mettent aussi en évidence une part de variance expliquée par les connaissances morphologiques qui constituent un prédicteur puissant des performances ultérieures en lecture. En revanche, les capacités de dénomination rapide ne semblent pas contribuer à l'explication de la variance des performances en lecture en langue arabe. Les évaluations des enseignants jouent également un rôle important dans la prédiction de la réussite dans l'apprentissage de la lecture et dans le repérage des difficultés des élèves.

L'étude de la valeur prédictive individuelle montre que les évaluations langagières et cognitives et les évaluations réalisées par les enseignants prédisent mieux les futurs bons lecteurs que ceux qui éprouveront des difficultés dans ce domaine (sensibilité et valeur prédictive positive faibles).

MOTS CLÉS : Apprentissage de la lecture, langue arabe, étude prédictive, compétences langagières, compétences cognitives, identification de mots, diglossie, évaluation par les enseignants, valeur prédictive individuelle, prévention, modèle d'équations structurelles.

TITLE : Longitudinal and predictive study of success or failure in learning to read in Arabic

ABSTRACT

This longitudinal study aims to explore the main overall and individual predictive value of early language and cognitive skills on performance in reading in Arabic. Secondly, the aim is to determine the ability of an evaluation of behavior and academic skills informed by classroom teachers preparatory (CP) to predict the next level in reading. The results obtained from 171 Tunisian children followed the preparatory class in the 2nd year of primary school confirm the predictive value of letter knowledge and phonological knowledge. They also highlight a portion of variance explained by the morphological knowledge as a powerful predictor of subsequent performance in reading. In contrast, the capacity of rapid naming did not seem to contribute to the explanation of the variance in performance in reading in Arabic. Teachers rating scale also play an important role in predicting success in learning to read and tracking of student difficulties. The study of the individual predictive value shows that language and cognitive assessments and evaluations by teachers are better predictors of future good readers as those who struggle in this domain (low sensitivity and positive predictive value).

KEY WORDS : Learning to read, Arabic language, predictive study, language skills, cognitive skills, words identification, diglossia, teacher assessment, individual predictive value, prevention, structural equation model.

DISCIPLINE : Psychologie

Université de Nantes
Laboratoire de Psychologie "Education, Cognition, Développement" (EA 3259)
Chemin de la Censive du Tertre – BP 81 227 – 44312 NANTES Cedex 3