

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LA MÉMOIRE À COURT TERME ET LE VOCABULAIRE COMME
PRÉDICTEURS DES PERFORMANCES EN DÉCODAGE ET EN
COMPRÉHENSION DE LA LECTURE :
UNE ANALYSE LONGITUDINALE CHEZ LES ENFANTS INNUS

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN LINGUISTIQUE

PAR
ROSALIE BOURDAGES

NOVEMBRE 2014

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens d'abord à remercier ma directrice, Lori Morris, qui a su me guider tout au long de la rédaction de ce mémoire. En me donnant la possibilité de collaborer à son projet dans la communauté de Pessamit, elle m'a permis de confirmer mon intérêt pour la linguistique appliquée et le travail de terrain. Je remercie aussi mon co-directeur, Denis Foucambert, pour son aide constante, en particulier en ce qui a trait aux analyses statistiques. Je suis aussi extrêmement reconnaissante envers mes lectrices, Marie Labelle et Véronique Fortier, pour les précieux conseils qu'elles m'ont prodigués.

Je souhaite finalement exprimer toute ma gratitude aux élèves, aux enseignants et à l'ensemble du personnel de l'école primaire Nussim à Pessamit.

TABLE DES MATIÈRES

| | |
|---|-----|
| REMERCIEMENTS..... | i |
| LISTE DES FIGURES | iii |
| LISTE DES TABLEAUX | iii |
| LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES..... | iii |
| RÉSUMÉ | ix |
| INTRODUCTION | 1 |
| CHAPITRE I | |
| PROBLÉMATIQUE..... | 5 |
| 1.1 Questions de recherche | 7 |
| CHAPITRE II | |
| CADRE THÉORIQUE..... | 10 |
| 2.1 La lecture | 10 |
| 2.1.1 Le décodage..... | 13 |
| 2.1.2 La compréhension en lecture..... | 16 |
| 2.2 Les mémoires..... | 18 |
| 2.2.1 La mémoire de travail et la mémoire à court terme | 19 |
| 2.2.2 Retour sur la mémoire à court terme | 23 |
| 2.2.3 Les mémoires en lien avec la lecture..... | 25 |
| 2.3 Le vocabulaire..... | 27 |
| 2.3.1 Le vocabulaire et les propriétés des mots..... | 27 |
| 2.3.2. Le vocabulaire en lien avec la lecture | 30 |

CHAPITRE III

| | |
|--|----|
| MÉTHODE..... | 34 |
| 3.1 Description du corpus | 34 |
| 3.1.1 Collectes de données | 35 |
| 3.1.2 Participants et exclusions | 36 |
| 3.2 Description des épreuves des temps 1, 2 et 3 | 37 |
| 3.2.1 Épreuves mesurant le vocabulaire..... | 38 |
| 3.2.2 Épreuves mesurant la mémoire à court terme | 42 |
| 3.3 Description des épreuves du temps final | 46 |
| 3.3.1 Épreuve mesurant la compréhension en lecture | 46 |
| 3.3.2 Épreuves mesurant le décodage | 51 |
| 3.3.3 Mise à l'essai..... | 53 |
| 3.4 Synthèse des variables et des épreuves..... | 54 |
| 3.5 Traitement des données | 57 |

CHAPITRE IV

| | |
|--|----|
| RÉSULTATS..... | 59 |
| 4.1 Analyses appliquées aux variables mesurées..... | 59 |
| 4.1.1 Statistiques descriptives pour les variables mesurées | 59 |
| 4.1.2 Analyses de corrélation | 62 |
| 4.1.3 Analyses de régression | 68 |
| 4.2 Analyses appliquées aux habiletés latentes | 76 |
| 4.2.1 ACP permettant la construction des variables latentes | 76 |
| 4.2.2 Statistiques descriptives des variables latentes construites par ACP | 81 |
| 4.2.3 Analyses de corrélation des variables latentes construites par ACP | 81 |
| 4.2.4 Analyses de régression et analyses en composantes principales (ACP) | 83 |

CHAPITRE V

| | |
|--|----|
| DISCUSSION..... | 88 |
| 5.1 Réponse à la première question de recherche..... | 88 |

| | |
|--|-----|
| 5.1.1 Épreuves permettant de prédire les performances en compréhension | 89 |
| 5.1.2 Épreuves permettant de prédire les performances en décodage | 91 |
| 5.2 Réponse à la deuxième question de recherche..... | 92 |
| 5.2.1 Le vocabulaire en lien avec la compréhension en lecture | 93 |
| 5.2.2 Le vocabulaire en lien avec le décodage | 94 |
| 5.2.3 La MCT en lien avec la compréhension en lecture | 95 |
| 5.2.4 La MCT en lien avec le décodage | 97 |
| 5.3 Considérations d'ordre méthodologique..... | 98 |
| 5.3.1 Considérations méthodologiques particulières..... | 98 |
| 5.3.2 Considérations méthodologiques générales | 99 |
| CONCLUSION..... | 101 |
| ANNEXE A | |
| ÉPREUVES DU TEMPS FINAL..... | 104 |
| ANNEXE A.1 | |
| ÉPREUVE DE LECTURE DE MOTS EN UNE MINUTE | 105 |
| ANNEXE A.2 | |
| ÉPREUVE DE LECTURE DE NON-MOTS EN UNE MINUTE..... | 106 |
| ANNEXE A.3 | |
| ÉPREUVE DE COMPRÉHENSION EN LECTURE..... | 107 |
| ANNEXE B | |
| SUPPLÉMENTS AUX ANALYSES..... | 112 |
| ANNEXE B.0 | |
| Comparaison des résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence..... | 113 |
| ANNEXE B.1 | |
| Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 1 | 115 |

ANNEXE B.2

| | |
|--|-----|
| Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 2 | 116 |
|--|-----|

ANNEXE B.3

| | |
|--|-----|
| Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 3 | 117 |
|--|-----|

| | |
|--------------------|-----|
| BIBLIOGRAPHIE..... | 118 |
|--------------------|-----|

LISTE DES FIGURES

| | |
|--|----|
| Figure 1.1 – Modélisation de la deuxième question de recherche ainsi que des deux sous-questions, tous temps confondus | 9 |
| Figure 2.1 – Architecture du modèle de lecture DRC de Coltheart <i>et al.</i> (2001), tiré de (Maionchi-Pino, 2008)..... | 15 |
| Figure 2.2 – Modèle de la mémoire de travail (Baddeley, 2000) – inspiré de Fortier (2013)..... | 21 |
| Figure 2.3 – La structure de l'entrée lexicale – adapté de Levelt (1989) (cité dans Jiang, 2000)..... | 28 |
| Figure 4.1 – Diagramme de composantes (temps 1) | 78 |
| Figure 4.2 – Diagramme de composantes (temps 2) | 79 |
| Figure 4.3 – Diagramme de composantes (temps 3) | 80 |
| Figure 4.4 – Modélisation des résultats pour les analyses de régression effectuées avec les variables latentes | 87 |

LISTE DES TABLEAUX

| | |
|---|----|
| Tableau 2.1 – Comparaison des caractéristiques de la MdT et de la MCT | 25 |
| Tableau 3.1 – Portrait de la fréquence du vocabulaire utilisé dans l'épreuve de compréhension en lecture selon la base <i>Lextutor</i> (Cobb, 2013) | 48 |
| Tableau 3.2 – Pointage accumulé pour les bonnes réponses (notées sur 2) pour chaque texte de l'épreuve de compréhension en lecture (n=71) | 50 |
| Tableau 3.3 – Synthèse des habiletés latentes anticipées et des épreuves | 55 |
| Tableau 4.1 – Statistiques descriptives des variables à l'étude retenues pour l'analyse (après transformations) | 61 |
| Tableau 4.2 – Corrélations entre les variables du T1 et du TF | 63 |
| Tableau 4.3 – Corrélations entre toutes les variables (Temps 2) | 66 |
| Tableau 4.4 – Corrélations entre toutes les variables (Temps 3) | 67 |
| Tableau 4.5 – Six modèles de régression testés (après ajustements) | 70 |
| Tableau 4.6 – Résultats pour les analyses de régression pour le décodage au temps 3 | 72 |
| Tableau 4.7 – Résultats pour les analyses de régression pour la compréhension au temps 1 | 73 |
| Tableau 4.8 – Résultats pour les analyses de régression pour la compréhension au temps 2 | 74 |
| Tableau 4.9 – Résultats pour les analyses de régression pour la compréhension au temps 3 | 75 |
| Tableau 4.10 – Statistiques relatives aux composantes créées par ACP aux temps 1, 2 et 3 | 77 |
| Tableau 4.11 – Statistiques descriptives des variables latentes à l'étude retenues pour l'analyse | 82 |
| Tableau 4.12 – Corrélations entre toutes les variables (Temps 1) | 82 |

| | |
|---|----|
| Tableau 4.13 –Six modèles de régression testés..... | 85 |
| Tableau 4.14 Résultats pour les analyses de régression pour la compréhension au temps 1..... | 86 |
| Tableau 4.15 – Résultats pour les analyses de régression pour la compréhension au temps 3..... | 86 |
| Tableau 5.1 – Synthèse des résultats sur les épreuves des T1, T2 et T3 permettant de prédire les performances en décodage et en compréhension au TF | 89 |

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

| | |
|------|--|
| ACP | Analyses en composantes principales |
| L1 | Langue maternelle |
| L2 | Langue seconde |
| MCT | Mémoire à court terme |
| MdT | Mémoire de travail |
| MLT | Mémoire à long terme |
| MSL | Modèle simple de lecture |
| NÉEL | <i>Nouvelles épreuves pour l'examen du langage</i> (batterie d'évaluation du langage oral) |
| T1 | Temps 1 (maternelle 5 ans, automne) |
| T2 | Temps 2 (maternelle 6 ans, printemps) |
| T3 | Temps 3 (première année, printemps) |
| TF | Temps final (3 ^e et 4 ^e année, printemps) |
| Voc | Vocabulaire |

RÉSUMÉ

Les enfants des communautés autochtones du Canada présentent des lacunes par rapport aux moyennes québécoise et canadienne sur le plan de la lecture (CMEC, 2008). Or la maîtrise de la langue écrite est garante de la réussite à l'école et de l'intégration sociale et professionnelle (Hemphill et Tivnan, 2008). La lecture, traditionnellement représentée comme l'interaction entre le décodage et la compréhension (Gough et Tunmer, 1986), s'appuie sur plusieurs habiletés psycholinguistiques, parmi lesquelles la mémoire à court terme (Alloway et Gathercole, 2005) et le vocabulaire (Lervåg et Aukrust, 2010).

À l'aide de données longitudinales récoltées à quatre moments successifs auprès d'enfants innus ($n=75$), notre étude cherche à modéliser et à évaluer le rôle prédictif que jouent ces habiletés dans les performances ultérieures en décodage et en compréhension. Les participants ont passé trois épreuves de vocabulaire (d'expression et de réception) et quatre épreuves associées à la mémoire à court terme (répétition de mots rares, de chiffres, de phrases et de rythmes).

D'une part, ce travail permet de comprendre, à l'aide d'analyses en composantes principales, comment les différentes épreuves construisent deux habiletés latentes homogènes qui correspondent au vocabulaire et à la mémoire à court terme. D'autre part, les analyses de régression permettent d'évaluer la contribution de ces habiletés, soit la mémoire à court terme et le vocabulaire, aux performances en lecture. La capacité différenciée de chacune des épreuves à prédire les performances ultérieures en lecture est aussi évaluée. Les résultats illustrent la relation étroite qu'entretiennent le vocabulaire et la compréhension en lecture chez les jeunes Innus.

MOTS-CLÉS: mémoire à court terme, vocabulaire, compréhension en lecture, décodage, habileté cognitive, habileté langagière, enfants d'une communauté innue.

INTRODUCTION

Le passage par l'école primaire permet aux élèves de développer des habiletés et des compétences en littératie, comprise comme la compréhension et l'utilisation des informations écrites, qui les incitent à poursuivre leur cheminement scolaire et à éventuellement contribuer pleinement à la société (Desrochers, Carson et Daigle, 2012). Dans un parcours scolaire québécois typique, la lecture fait l'objet d'un enseignement explicite au début du primaire, puis est utilisée comme un outil permettant l'acquisition de connaissances dans toutes les matières enseignées. Elle constitue, en quelque sorte, la « colonne vertébrale des apprentissages » (Bertrand, Fluss, Billard et Ziegler, 2010, p. 300). Les performances en lecture sont un gage attesté de la réussite scolaire et un tremplin vers une intégration sociale et professionnelle (Hemphill et Tivnan, 2008 ; OCDE, 2012 ; Snow, Burns et Griffin, 1998).

Snow *et al.* (1998, p. 20) résumant bien les facteurs qui peuvent influencer les performances en lecture et présentent trois obstacles qui peuvent freiner le développement d'habitudes de lecture :

There are three potential stumbling blocks that are known to throw children off course on the journey to skilled reading. The first obstacle, which arises at the outset of reading acquisition, is difficulty understanding and using the alphabetic principle – the idea that written spellings systematically represent spoken words. It is hard to comprehend connected text if word recognition is inaccurate or laborious. The second obstacle is a failure to transfer the comprehension skills of spoken language to reading and to acquire new strategies that may be specifically needed for reading. The third obstacle to reading will magnify the first two: the absence or loss of an initial motivation to read or failure to develop a mature appreciation

of the rewards of reading.

Les deux premières barrières renvoient en fait à deux des éléments dont la lecture est constituée : la correspondance graphème-phonème – c'est-à-dire le décodage – et la compréhension. Cela implique que des lacunes dans l'une ou l'autre de ces composantes peuvent hypothéquer les performances en lecture (Gough et Tunmer, 1986). La troisième barrière constitue les conditions dans lesquelles évoluent les lecteurs et les lectrices ; les auteurs mettent l'accent sur la réalité sociale et personnelle, qui est susceptible d'influencer les performances en lecture. Ainsi, un environnement équilibré qui suscite un intérêt général pour l'activité de lecture se compose d'un enseignement explicite de la lecture, d'un milieu scolaire stimulant et d'un nombre réduit de facteurs de risque, comme la pauvreté et l'exclusion (Badian, 1990).

Nous avons fait le choix d'étudier les éléments constitutifs de la lecture, soit le décodage et la compréhension en lecture, dans une communauté autochtone qui vit, comme d'autres au Canada, d'importants problèmes nuisant à la réussite scolaire de ses jeunes. Globalement, les jeunes autochtones font face à beaucoup plus d'obstacles pendant leur parcours scolaire que les jeunes non autochtones, et ce, dans un contexte où les ressources sont limitées. En effet, on observe dans les communautés des Premières Nations une faiblesse dans la maîtrise du vocabulaire par rapport aux moyennes canadiennes, des problèmes de mémoire et d'attention, un manque de confiance et un manque de motivation envers l'école (CMEC, 2008 ; Findley, Kohen et Miller, 2014). Les enfants autochtones sont aussi plus à risque de souffrir du syndrome d'alcoolisation fœtale, parmi d'autres conséquences de l'abus de substances (April et Bourret, 2004).

Dans la communauté innue de Pessamit, à laquelle nous nous intéresserons particulièrement, une multitude de facteurs sont à la source des problèmes scolaires des élèves. Parmi ceux-ci, on peut citer la pauvreté, le bas niveau d'éducation des parents et le contexte diglossique dans lequel évoluent les enfants, certains s'exprimant encore dans la langue ancestrale (Morris, 2013 ; Morris et Mackenzie, 2012 ; Oudin et Drapeau, 1993). Nombreuses sont les études qui associent un tel portrait avec la présence de lacunes en lecture chez les enfants, et ce, tôt dans leur développement, ce qui compromet leur acquisition de la littératie (Hemphill et Tivnan, 2008 ; Snow *et al.*, 1998 ; Whitehurst, 1997). La notion d'élèves « à risque » renvoie à une catégorie de jeunes qui éprouvent des difficultés scolaires et dont le passage à la vie adulte et l'entière contribution à la société sont compromis (Lurin et Soussi, 2013). En effet, les difficultés scolaires et sociales que vivent les élèves de Pessamit peuvent mener au décrochage scolaire (Fortin et Picard, 1999). De plus, certains enfants de la communauté commencent leur scolarisation en français avec des lacunes en conscience grapho-phonologique, ce qui risque d'affecter le développement d'habiletés en lecture (Bodson, 2013).

Le premier objectif que nous poursuivons est de nature diagnostique. Nous souhaitons déterminer quelles épreuves, parmi celles soumises en trois temps distincts à des élèves autochtones, nous permettront de mieux prédire les performances en décodage et en compréhension en un temps subséquent. Le dépistage précoce et efficace des problèmes pouvant affecter la maîtrise ultérieure de la lecture chez certains enfants revêt une importance particulière dans les communautés étant confrontées à des problèmes socio-économiques difficiles. Nous souhaitons que notre recherche ait des implications préventives et puisse contribuer à freiner la progression des problèmes que vit la communauté.

Notre deuxième objectif est de nature explicative. Nous nous appliquerons à comprendre comment s'articulent, dans un contexte particulier, les liens entre le vocabulaire, la mémoire et les composantes de la lecture.

Dans le premier chapitre de ce mémoire, nous échafauderons notre problématique de recherche. Le deuxième chapitre constitue une revue de la littérature sur la mémoire et sur le vocabulaire, toujours en lien avec le développement des éléments compositionnels de la lecture. La présentation de la méthode utilisée fera l'objet du troisième chapitre. Dans le quatrième chapitre, nous ferons un compte-rendu des résultats des analyses statistiques effectuées, pour poursuivre, au cinquième chapitre, avec une discussion de ces résultats. Nous concluons avec des considérations d'ordres théorique et diagnostique quant à la problématique étudiée.

CHAPITRE I

PROBLÉMATIQUE

Le dépistage des difficultés scolaires chez les Innus n'a pas fait l'objet de beaucoup de recherches, la communauté scientifique ne s'étant pas suffisamment penchée sur la spécificité de ces populations jusqu'à présent. Nous savons pourtant que le parcours scolaire des élèves autochtones est cahoteux : ceux-ci présentent un retard par rapport aux normes canadiennes sur le plan de l'écriture et de la lecture et sont sujets à un décrochage scolaire prématuré (Morris et Mackenzie, 2013). Les résultats préliminaires de Morris (2014) semblent montrer que les enfants innus de Pessamit présentent des lacunes sur le plan du vocabulaire en réception et en expression dans la langue majoritaire – le français –, par rapport aux jeunes francophones du même âge. Ce domaine constitue probablement la plus grande faiblesse de ces élèves (Morris, 2014 ; Morris et Labelle, 2008). La capacité à répéter des chiffres et des mots, qui constituent des mesures classiques de mémoire verbale à court terme (Conway, Kane, Bunting, Hambrick, Wilhelm et Engle, 2005 ; Engle, Laughlin, Tuholski et Conway, 1999), présentait aussi des lacunes.

Notre analyse des données longitudinales recueillies auprès d'enfants de Pessamit par Morris et ses collaboratrices entre 2008 et 2012 permet d'identifier les épreuves qui prédisent le mieux la performance en lecture. Nous souhaitons que l'analyse contenue dans ce travail de recherche ait des implications pédagogiques et préventives.

Certaines des épreuves qui ont été soumises chaque année à ces enfants tendent à circonscrire les capacités cognitives associées à la mémoire à court terme, tandis que d'autres mesurent les connaissances lexicales. Nos considérations d'ordre diagnostique possèdent un indispensable corollaire : l'analyse du lien qui unit la mémoire à court terme et le vocabulaire avec les performances ultérieures en lecture. Les élèves de la communauté à l'étude sont à risque puisqu'ils vivent à l'intérieur d'un milieu socialement défavorisé, ce qui les prédispose à rencontrer des difficultés scolaires. Nous cherchons à dépister des éléments spécifiques qui expliqueraient une partie de leurs difficultés.

Nous souhaitons donc examiner l'impact de ces éléments – mémoire à court terme et vocabulaire – sur le décodage et la compréhension, afin d'établir de quelle façon ils contribuent aux performances en lecture. En effet, la mémoire à court terme a été corrélée avec l'acquisition de la compréhension en lecture (Alloway et Gathercole, 2005 ; Martinez Perez, Majerus et Poncelet, 2012) et le développement langagier. La connaissance du vocabulaire constitue également un prédicteur essentiel du développement de la compréhension en lecture (Lervåg et Aukrust, 2010). Dans la mesure où ces habiletés ont souvent été mises en relation avec les performances en lecture, notre travail de recherche permettra de vérifier empiriquement ce lien et ce, à l'aide de données recueillies dans une communauté particulièrement vulnérable aux problèmes scolaires.

Pour mener à bien notre projet de recherche, nous feront appel à certains modèles psycholinguistiques qui ont permis de grandes avancées dans le domaine de l'acquisition et de la prédiction de la performance en lecture (Rastle, 2009). Ils ont permis de tester des hypothèses, de développer des théories générales sur ce qu'implique la lecture et d'ainsi mieux expliquer les problèmes en lecture.

La présente étude se penche sur la contribution de la mémoire à court terme, d'une part, opérationnalisée par les tâches de répétitions (phrases, mots rares, chiffres, rythmes) et, d'autre part, de l'étendue du vocabulaire en expression et en réception (concepts mathématiques, dénomination, topologie et arithmétique), à l'explication des différences observées chez les participants sur le plan des performances en lecture. L'apport principal de cette étude est de vérifier comment les liens reconnus entre le vocabulaire et les composantes de la lecture, puis entre la mémoire à court terme et les composantes de la lecture, se présentent chez les enfants d'une communauté innue.

1.1 Questions de recherche

Cette recherche s'inscrit dans un ensemble de recherches visant à mesurer l'influence d'habiletés langagières et cognitives à la performance en lecture, et ce, dans une approche longitudinale. Elle poursuit deux objectifs : dans un premier temps, l'évaluation de la validité de différentes épreuves isolées dans la prédiction des résultats en lecture; dans un deuxième temps, l'évaluation des compétences transversales qu'il faut mettre à contribution pour réussir les épreuves, en tenant compte de la spécificité de la communauté à l'étude. La complexité de la situation propre à la population à l'étude nécessite une analyse plus fine des compétences mises à contribution dans la réalisation des épreuves qui seront analysées. Cette étude est donc motivée par les questions suivantes :

- 1) Dans quelle mesure est-ce que chacune des épreuves soumises aux participants de la maternelle à la 2^e année du primaire – répétition de phrases, mots rares, chiffres et de rythmes, topologie et arithmétique, vocabulaire (expression) et vocabulaire (réception) – est prédictive des performances en lecture en 3^e et 4^e année?

Cette question est de nature diagnostique. Ainsi, nous supposons que les épreuves permettront de dépister les problèmes ultérieurs en lecture, puisqu'elles prédiront les performances en décodage et en compréhension. La question de recherche précédente a son pendant explicatif :

- 2) Comment les habiletés latentes que constituent le vocabulaire et la mémoire à court terme contribuent à expliquer les performances en lecture en 3^e et 4^e année du primaire?

Cette deuxième question de recherche requiert que nous répondions au préalable aux deux sous-questions suivantes :

- 3) Dans quelle mesure est-ce que les épreuves de répétition de mots rares, de phrases, de chiffres et de rythmes, soumises aux participants entre la maternelle et la 2^e année du primaire, construisent une variable latente qui représente la mémoire à court terme?
- 4) Dans quelle mesure est-ce que les épreuves de topologie et arithmétique, de vocabulaire (expression) et de vocabulaire (réception), soumises aux participants entre la maternelle et la 2^e année du primaire, construisent une variable latente qui représente les habiletés en vocabulaire?

Ainsi, nous supposons, d'une part, que les épreuves de répétition de mots rares, de chiffres, de phrases et de rythmes circonscrivent une habileté assimilée à la mémoire à court terme. D'autre part, nous considérons que les épreuves de topologie et arithmétique, de vocabulaire (expression) et de vocabulaire (réception) circonscrivent une habileté assimilée au vocabulaire. Au vu des études antérieures, nous faisons l'hypothèse que ces habiletés, telles que construites par nos épreuves, contribueront à expliquer les performances ultérieures en décodage, telles que mesurées à l'aide des

épreuves de lecture de mots et de non-mots, et en compréhension en lecture, telle que mesurée par l'épreuve de compréhension de textes courts. Les hypothèses de recherche sont résumées par le modèle suivant, où les habiletés latentes sont en gris clair (figure 1.1).

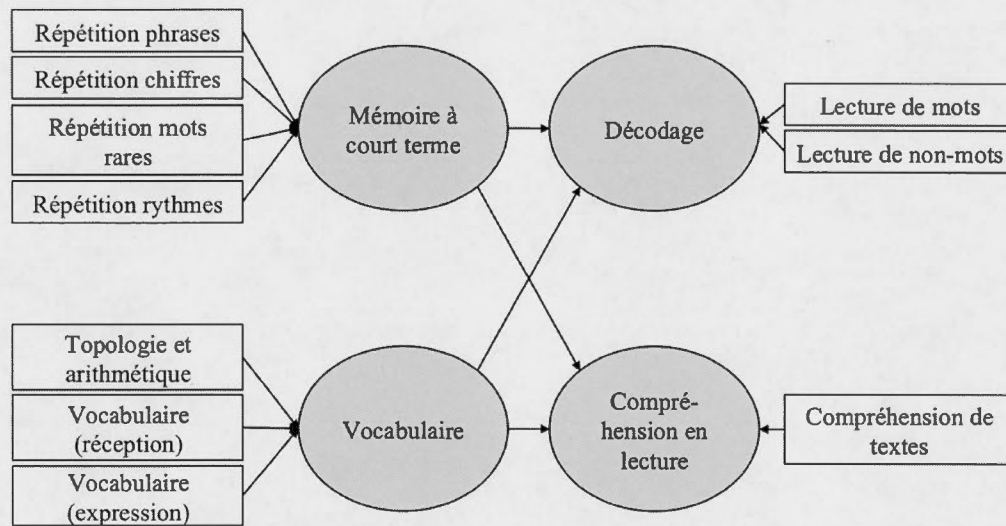


Figure 1.1 – Modélisation de la deuxième question de recherche ainsi que des deux sous-questions, tous temps confondus

Le chapitre suivant présente les travaux théoriques et expérimentaux qui nous permettront d'étayer nos questions de recherche.

CHAPITRE II

CADRE THÉORIQUE

Dans ce chapitre, nous passerons en revue les travaux sur la lecture ainsi que ceux se consacrant à la mémoire de travail et la mémoire à court terme, puis nous terminerons avec une recension des écrits sur le vocabulaire.

2.1 La lecture

Perfetti (1989, p. 62) définit l'acte de lire comme « un ensemble de processus permettant d'extraire la signification du texte ». L'ampleur et la complexité de la tâche qui est demandée au lecteur ou à la lectrice résident justement dans les processus qui sous-tendent cet accès au sens. Le lecteur habile a automatisé ces processus. Il lit avec aisance, c'est-à-dire rapidement, sans effort et avec exactitude (Rastle, 2009).

La lecture constitue un mécanisme coûteux sur le plan cognitif : elle fait appel à plusieurs capacités de différents niveaux, dont la mémoire à court terme, la mémoire de travail, la mémoire à long terme, les connaissances lexicales, l'accès au lexique, les processus syntaxiques, la compréhension et d'autres habiletés encore (Hoover et Tunmer, 1993 ; Joshi et Aaron, 2000). Jusqu'à la fin du 20^e siècle, c'est d'ailleurs le modèle complexe de la lecture (*Complex View of Reading*) qui a dominé la littérature. Ce modèle considère la lecture comme un regroupement diversifié et complexe de processus mentaux participant de toute forme de pensée – raisonnement, résolution de problèmes, jugement (Hoover et Gough, 1990 ; Hoover et Tunmer, 1993).

L'approche qui domine actuellement, particulièrement dans le cadre de la littérature se consacrant au diagnostic des dysfonctionnements en lecture, se nomme le modèle componentiel de la lecture – *Component Model of Reading*. Ce modèle est basé sur la *Simple View of Reading*, telle que définie dans les travaux classiques de Perfetti (1977), de Gough et Tunmer (1986) et de Hoover et Gough (1990). Bien qu'il s'inscrive en rupture avec le modèle complexe de la lecture, le modèle simple de lecture (dorénavant MSL) reconnaît la complexité des mécanismes opérant lors de la lecture ; il en propose toutefois une modélisation qui s'avère plus opératoire (Hoover et Tunmer, 1993). Ainsi, la compréhension en lecture est représentée comme l'interaction entre deux composantes : le décodage et la compréhension linguistique (Ecalte, 1997 ; Hoover et Gough, 1990 ; Hoover et Tunmer, 1993 ; Magnusson et Nauclér, 1991 ; Perfetti, 1989). Ces deux composantes sont considérées comme théoriquement distinctes, empiriquement isolables, indispensables et, selon Hoover et Tunmer (1993) et Hoover et Gough (1990), d'importance égale. Ainsi, aucune des deux composantes ne suffit, à elle seule, à assurer les performances optimales en lecture. La dyslexie, phénomène par lequel des habiletés fragmentaires en décodage coexistent avec une compréhension linguistique intacte, démontre la divisibilité de ces deux construits théoriques (Hoover et Gough, 1990). Ces assertions sont traditionnellement représentées par une équation élémentaire, où la compréhension en lecture (L) est le produit du décodage (D) et de la compréhension linguistique (C) – $L = D \times C$. Chez Perfetti (1977) (cité dans Hoover et Tunmer, 1993), on retrouve les mêmes composantes, mais dans une relation d'addition, où $L = D + C$. La composante C, soit la compréhension linguistique, est comprise comme l'habileté à générer des informations au niveau phrastique ou textuel sur la base d'informations orales de niveau lexical (Gough et Tunmer, 1986 ; Hoover et Gough, 1990 ; Hoover et Tunmer, 1993 ; Tunmer et Chapman, 2012). Ces deux équations, issues d'une analyse componentielle de la lecture, souscrivent donc à l'hypothèse forte du MSL, selon laquelle la variance en lecture peut être expliquée par deux grands groupes d'habiletés, soit le décodage et la compréhension linguistique. La première formule suggère que

chacune des deux composantes sont susceptibles d'interagir l'une avec l'autre (Hoover et Gough, 1990 ; Lecocq, Casalis, Leuwers et Watteau, 1996, p. 109). Le choix de l'une ou l'autre de ces combinaisons ne se pose toutefois pas dans le cadre de ce mémoire.

Le décodage étant une habileté typiquement de bas niveau, plusieurs études se sont donné pour mandat d'évaluer la contribution de cette variable à la compréhension en lecture ou à la production d'inférences (Cain, Oakhill et Bryant, 2004a ; Lervåg et Aukrust, 2010). La littérature considère massivement qu'un bon décodeur deviendra un bon lecteur, c'est-à-dire que les habiletés en décodage constituent un indicateur fiable du niveau ultérieur en compréhension en lecture, et que la maîtrise du décodage est un préalable indispensable à une lecture efficace (Gough et Tunmer, 1986 ; Snow *et al.*, 1998 ; Tunmer et Chapman, 2012). Les résultats de Lervåg et Aukrust (2010), nuancent cette assertion. En effet, les auteurs ont observé une contribution non-significative des habiletés précoces en décodage sur le développement de la compréhension en lecture mesurée 18 mois plus tard, et ce, tant chez les apprenants L1 que L2. Les résultats de Magnusson et Naclér (1991) suggèrent également de renverser la directionnalité du lien décodage-compréhension : ce sont les lecteurs qui comprennent bien le texte qui deviennent de bons décodeurs (lisent vite et en faisant peu d'erreurs). La représentation sérielle de la lecture va jusqu'à influencer la façon dont elle est enseignée en classe : d'abord on expose aux élèves les conventions de la correspondance graphème-phonème, ensuite vient la compréhension des textes. La lecture est réputée devenir un moyen d'apprendre : l'interprétation prend le pas sur le décodage (Desrochers *et al.*, 2012 ; Magnusson et Naclér, 1991).

Ainsi, la séquentialité des deux processus est matière à débat. Ce débat outrepassé toutefois le dessein de ce mémoire, où l'on s'attarde aux performances en lecture – comprises comme les performances en décodage (D) et en compréhension (L). Nous nous inspirerons toutefois de cette représentation des processus de lecture, où le

décodage – une habileté de bas niveau – participe de la compréhension en lecture – une habileté de haut niveau. La lecture est ainsi envisagée, dans le cadre de ce mémoire, de façon statique, comme un produit fini ; l'enchaînement des processus mis en œuvre pour arriver à ce résultat final n'est pas notre propos. Le décodage et la compréhension en lecture seront considérés comme les deux habiletés finales qui nous permettront de statuer sur la performance en lecture des élèves que nous étudierons, le tout sous l'égide du MSL. Les sections suivantes sont dédiées à la définition de chacune de ces habiletés.

2.1.1 Le décodage

Le décodage se réfère à la capacité à transformer un input graphique en un code phonémique (Perfetti, 1985). Cette composante a été définie par Gough et Tunmer (1986) comme l'habileté, subordonnée à la connaissance des correspondances graphèmes-phonèmes, permettant de lire des mots rapidement, silencieusement et avec précision. Le décodage habile se caractérise par son automaticité.

Alors que l'enfant apprend sa langue de façon naturelle, par exposition, le décodage du code écrit doit faire l'objet d'un apprentissage explicite (Desrochers et Berger, 2011). Le principe alphabétique constitue une notion incontournable dans la littérature sur le décodage étant donné l'abondance de travaux effectués avec des populations utilisant une langue alphabétique (comme l'anglais ou le français). Ce principe correspond à la capacité à se représenter un phonème à partir de combinaisons de lettres (Aarnoutse, Van Leeuwe, Voeten et Oud, 2001). Les enfants appartenant à notre communauté à l'étude, à l'instar de la plupart des élèves québécois, apprennent à décoder, c'est-à-dire faire puis automatiser les correspondances entre les graphèmes et les phonèmes, dès la maternelle, à 5 ans.

C'est la réussite à ce processus qui permet l'identification de séquences de graphèmes lues pour la première fois, comme les mots nouveaux et les mots rares. Ce sont sur ces bases que les représentations orthographiques lexicales sont elles-mêmes automatisées et stockées en mémoire (Ehri, 2005). L'importance du décodage dans la variation des performances en lecture est réputée décroître avec l'âge : elle est prépondérante chez les enfants qui commencent à lire, et moindre chez les enfants plus vieux (Aarnoutse *et al.*, 2001). Cette idée d'un effet moins important du décodage à long terme est cohérente avec le MSL (Hoover et Gough, 1990 ; cité dans Lervåg et Aukrust, 2010).

La modélisation de l'accès au lexique impliquant le décodage a fait l'objet de plusieurs travaux dans les dernières années, dans le domaine de la psychologie cognitive. Les modèles se sont appuyés principalement sur des tâches de décision lexicale et de catégorisation sémantique (Rastle, 2009). Parmi les modèles d'accès au lexique qui ont marqué la littérature sur l'apprentissage de la lecture, on retrouve le *Dual Route Cascaded Model* (modèle DRC) (Coltheart, Rastle, Perry, Langdon et Ziegler, 2001). Celui-ci constitue le modèle explicatif le plus opérationnel, vu sa propension à faire le lien entre les représentations phonologiques et sémantiques et la reconnaissance des mots écrits, en plus de permettre la simulation informatique (Rastle, 2009). Selon ce modèle, le lecteur possède deux chemins ou voies pour accéder à la représentation du mot écrit : l'une est directe et passe par un traitement orthographique du mot, l'autre est indirecte et opère par assemblage sur la base des correspondances graphème-phonème (Coltheart *et al.*, 2001). La voie lexicale ou directe traite mots écrits, graphèmes de façon globale, tandis que dans la voie pré-lexicale, non lexicale ou indirecte, les lettres sont traitées une à une pour être recodées en phonèmes (voir la figure 2.1). Ce modèle fait donc intervenir explicitement le décodage au cours de la lecture de mots n'ayant pas encore d'entrée dans le lexique mental : le lecteur habile peut utiliser les deux routes, alors que le lecteur débutant est contraint d'utiliser la voie indirecte (Treimen et Kessler, 2009). La modélisation de la lecture est ainsi au moins

en partie construite par le décodage, de la même façon que les enfants doivent avoir des outils en décodage pour développer une performance efficace en lecture.

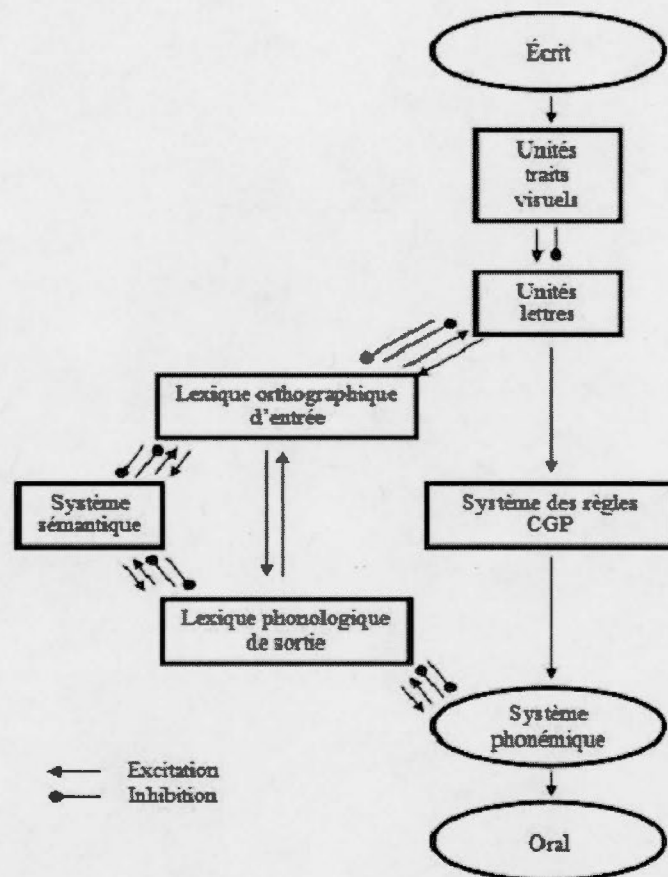


Figure 2.1 – Architecture du modèle de lecture DRC de Coltheart *et al.* (2001), tiré de (Mañonchi-Pino, 2008)

Par ailleurs, de nombreux travaux font état de la grande implication de la conscience phonologique dans les processus de lecture (Parrila, Kirby et McQuarrie, 2004). Certains suggèrent que le lien entre la mémoire à court terme et les performances en lecture passe par une médiation de la conscience phonologique (Baddeley, 1986 ;

Martinez Perez *et al.*, 2012). Il est effectivement raisonnable de penser qu'une segmentation appropriée de la chaîne parlée est indispensable à la maîtrise de la correspondance graphème-phonème (Tunmer, 1989). Le rôle de la conscience phonologique ne fait toutefois pas l'objet d'une réflexion profonde dans le cadre de ce travail. Bodson (2013) a recueilli des observations fort pertinentes sur la progression de cette habileté auprès des enfants de la communauté de Pessamit. Malheureusement, seuls les enfants de la maternelle de la communauté à l'étude ont participé aux tests de conscience phonologique, et les résultats se sont avérés peu concluants. En effet, la majorité des enfants n'étaient pas capables de réaliser les tâches demandées (Bodson, 2013).

2.1.2 La compréhension en lecture

La compréhension en lecture fait appel à un ensemble de processus de bas et de haut niveau. L'étude de la compréhension et des mécanismes qui la sous-tendent font toujours l'objet de dissensions au sein des sciences cognitives (Gernsbacher, 1991). Notre intention n'est pas de présenter le fonctionnement de l'apprentissage de la compréhension en lecture (Rémond et Quet, 1999), mais bien de décrire la performance en compréhension en lecture telle que nous la concevons à l'intérieur de ce mémoire, soit le succès de l'acte de comprendre un texte écrit.

Le lecteur expérimenté s'attend à ce que le texte qu'il lit soit cohérent (Oakhill, 1988, p. 5). La cohérence constitue une propriété de la compréhension textuelle que le lecteur établit entre certains éléments du texte (Gernsbacher, 1991). Johnson-Laird (1983) a proposé la notion de modèle mental, qui permet de représenter le texte et le monde qui est dénoté par le texte : le lecteur se construit donc une représentation en référence à ce qui est lu. La notion de modèle mental est une notion centrale dans la compréhension de texte, pour représenter le texte et le monde dont parle le texte. Van

Dijk et Kintsch (1983) ont bâti un modèle, appelé modèle de situation, qui distingue trois niveaux de représentations, ou niveaux de compréhension. On retrouve la structure de surface du texte – la reconnaissance des éléments lexicaux et syntaxiques du texte –, la base propositionnelle du texte – les inférences à partir des éléments qui n'étaient pas explicites dans le texte – et le modèle de situation du texte, c'est-à-dire le modèle mental créé à partir des interactions entre la phrase et sa représentation, les sujets principal et secondaire et l'information textuelle, ainsi que la connaissance qu'a le lecteur du monde. Ainsi, le discours s'échafaude avec des informations qui ne sont pas directement connectées entre elles. Les informations qui manquent et qui permettent de rendre le discours cohérent sont fournies par le lecteur, qui s'appuie sur sa connaissance du contexte et des faits (Van Dijk et Kintsch, 1983). Un texte est donc fondamentalement dans l'implicite; la présentation de toutes les informations et des propositions nécessaires à une interprétation cohérente du texte est en fait une entité théorique (Van Dijk et Kintsch, 1983).

La précédente définition de la compréhension en lecture est cohérente avec la composante L du MSL, qui est définie comme suit par Hoover et Tunmer (1993, p. 8) (notre traduction) :

Tant pour la compréhension linguistique que pour la compréhension en lecture, le modèle simple suppose une compréhension soignée : une compréhension qui est destinée à extraire du sens à partir du matériel présenté, par opposition à une compréhension visant à ne présenter que des idées principales ou à ne rechercher que des détails particuliers. Une mesure de la compréhension linguistique doit évaluer la capacité du sujet à comprendre la langue [orale]. [...] Une mesure de la compréhension de la lecture doit évaluer la même habileté, mais à partir du matériau imprimé (par exemple, en évaluant la capacité à répondre à des questions à propos du contenu d'un récit écrit).

Cette conception de la compréhension s'appuie essentiellement sur la production d'inférences et sur les processus d'intégration pour assurer son fonctionnement.

L'intégration correspond aux liens faits entre les différentes clauses se succédant, alors que les inférences renvoient aux événements et états qui sont nécessaires à une cohérence plus générale du texte (Cain *et al.*, 2004a, p. 32). Elles constituent les informations que construit le lecteur ou la lectrice à partir des éléments se trouvant de façon explicite dans le texte (Campion et Rossi, 1999). Ce sont ces deux procédés d'intégration et d'inférence qui permettent la construction d'un modèle textuel cohérent (Van Dijk et Kintsch, 1983). La compréhension d'un texte est donc tributaire des connaissances du lecteur. Cette connaissance du monde, qui inclut tant des informations de la vie quotidienne que des expériences plus personnelles, est indispensable à la production d'inférences (Campion et Rossi, 1999). Plusieurs études concluent d'ailleurs que les enfants qui sont faibles en compréhension de la lecture sont aussi faibles dans la production d'inférences permettant une construction cohérente du sens d'un texte (Cain, Oakhill et Lemmon, 2004b).

2.2 Les mémoires

La mémoire à court terme est le sujet qui nous préoccupe dans la mesure où elle correspond à ce qui est mesuré par nos épreuves (voir chapitre III). Nous nous efforcerons de montrer l'apport plus ou moins grand de la mémoire à court terme (désormais MCT) et des épreuves lui étant associées à la prédiction des performances en compréhension de la lecture chez la population à l'étude. Notre revue de la littérature sur la MCT s'inscrit dans le cadre théorique de la mémoire de travail (désormais MdT) de Baddeley (1986). Delafoy et Ehrlich (1990) écrivent d'entrée de jeu, dans leur revue de la littérature sur la MdT, que « les recherches visant à caractériser les processus qui régissent les activités cognitives complexes postulent toutes l'existence d'une mémoire de travail ». Il s'ensuit que la compréhension de la lecture doit être examinée dans le cadre du système cognitif, dont la MdT fait partie intégrante, dans la mesure où cette activité mobilise de façon importante les ressources

cognitives disponibles. Nous nous soumettons à cette tradition et procéderons à une analyse comparée de la MCT et de la MdT.

Les différences théoriques entre MCT et MdT sont toujours sujettes à débat. Certains auteurs décrivent les deux entités théoriques comme similaires, d'autres proposent que l'un est un sous-ensemble de l'autre, et vice versa. Engle *et al.* (1999) se sont employés à dresser un portrait critique de l'ambiguïté entourant la conception de ces deux importants construits théoriques. À l'instar de Baddeley (1986), de Just et Carpenter (1992) et d'autres (Klapp, Marshburn et Lester, 1983 ; Seigneuric, Ehrlich, Oakhill et Yuill, 2000), nous considérons que la MdT ne peut être mesurée que par le biais d'une tâche requérant la simultanéité des fonctions de traitement et de stockage de l'information. Nous prenons donc ici position en adoptant une conception différenciée de la MCT et de la MdT, où les deux systèmes remplissent des fonctions distinctes et sont représentées par des épreuves tout aussi distinctes. Nous poursuivrons avec une analyse différenciée des liens que ces deux construits cognitifs entretiennent avec la compréhension en lecture et le décodage.

2.2.1 La mémoire de travail et la mémoire à court terme

Commençons donc par circonscrire l'objet Mémoire. Globalement, la mémoire constitue un système permettant d'encoder, stocker, traiter et réutiliser l'information (Gathercole, 2009). Dans les années 1960, plusieurs études ont été consacrées à expliquer les effets de primauté et de récence. Le modèle explicatif d'Atkinson et Shiffrin (1968) avait retenu l'attention de la communauté scientifique, notamment parce qu'il permettait de rendre compte de ces effets (Delafoy et Ehrlich, 1990). Ces phénomènes se réfèrent à deux différents constats. D'abord, on avait observé que lors d'un rappel d'éléments sous forme de liste, ce sont les premiers (effet de primauté) et les derniers éléments de la liste (effet de récence) dont les sujets se souviennent le

mieux. Lorsqu'un rappel est fourni après plusieurs secondes, ce sont surtout les informations en début de liste qui sont répétées correctement. Ainsi, on a attribué à la mémoire à long terme (désormais MLT) la responsabilité du rappel des premiers éléments de la liste après un certain laps de temps. Le rappel efficace des derniers éléments de la liste a été attribué à la MCT. Atkinson et Shiffrin (1968) sont, de fait, les auteurs de l'une des premières modélisations de la mémoire comprenant plusieurs sous-systèmes. Leur modèle comprend trois composantes : la mémoire sensorielle, la mémoire à court terme et la mémoire à long terme. La MCT, telle que modélisée alors, était analogue à la MdT d'aujourd'hui, étant donné qu'elle stockait l'information reçue de la mémoire à long terme et de la mémoire sensorielle (Atkinson et Shiffrin, 1968). Baddeley (2003, p. 190) décrit en ces termes la composante MCT d'alors :

The most influential two-component model was that of Atkinson and Shiffrin (1968), who proposed that information came in from the environment into a temporary short-term storage system which served as an antechamber to the more durable LTM. In their model, the temporary system also served as a working memory, a workspace necessary not only for long-term learning, but also for many other complex activities such as reasoning and comprehension.

Par la suite, les travaux se penchant sur la modélisation des processus ayant trait à la mémoire ont en effet conclu à la différenciation de la MCT par rapport à la MdT (Baddeley et Hitch, 1974 ; Daneman et Carpenter, 1980 ; Klapp *et al.*, 1983 ; Swanson et Howell Ashbaker, 2000) : « *WM tasks require the monitoring of multiple capacity pools [...], whereas STM tasks access a passive storage system that draws upon a common pool of resources [...]* » (Swanson et Berninger, 1996, p. 363). Baddeley et Hitch (1974) proposent que le système mnésique est plutôt constitué d'une interaction entre la MdT et la MLT.

Le modèle de la MdT de Baddeley (1986) est subdivisé en plusieurs sous-systèmes. La MdT est le système qui permet, à la fois, la mémorisation et le traitement

d'informations durant une brève période de temps (Baddeley, 1986 ; Gathercole, 2009 ; Just et Carpenter, 1992). Sa spécificité réside donc dans le fait de procéder au traitement de certaines informations tout en les maintenant emmagasinées – un processus actif qui se distingue des processus passifs propres à la MCT, où les informations sont uniquement stockées puis rappelées. La tâche d'empan de lecture de Daneman et Carpenter (1980), qui s'inscrit dans une vision capacitaire du modèle de Baddeley, correspond à une mesure classique de la MdT. L'épreuve requiert que le sujet lise et comprenne une suite de phrases tout en retenant les derniers mots de chaque phrase. Non seulement les mots doivent-ils être rappelés, mais le lecteur doit aussi les intégrer dans un schéma de la phrase, puisqu'il aura à répondre à une question de compréhension s'y rapportant. L'empan de lecture correspond au nombre de mots finaux rappelés correctement. Le maintien en mémoire de l'adresse postale d'une personne tout en écoutant les instructions pour s'y rendre est un exemple de l'utilisation de la mémoire de travail dans la vie quotidienne.

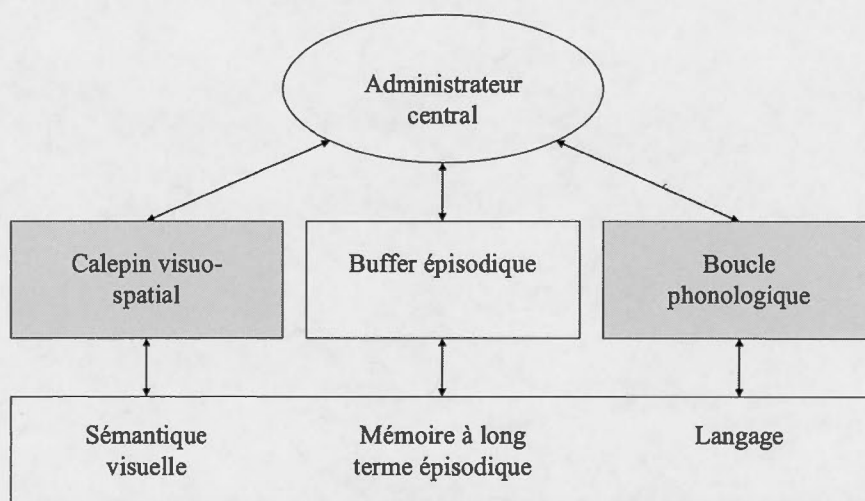


Figure 2.2 – Modèle de la mémoire de travail (Baddeley, 2000) – inspiré de Fortier (2013)

Note : Les composantes en gris clair constituent les systèmes esclaves de la MdT qui n'impliquent pas de traitement. Elles sont considérées, dans le cadre de ce mémoire, comme des entités appartenant à la MCT.

Le modèle de la MdT tel que proposé par Baddeley (Baddeley, 1986, 2000 ; Baddeley et Hitch, 1974) a subi plusieurs modifications depuis son introduction en 1974, mais s'avère néanmoins l'un des concepts théoriques les plus influents dans le domaine de la cognition. Son architecture est subdivisée en quatre sous-systèmes. L'administrateur central – *central executive* – constitue le sous-système autour duquel gravitent les autres construits. Il s'agit d'un système attentionnel qui contrôle et coordonne le traitement des informations fournies par les trois sous-systèmes, qui se chargent d'emmagasiner différents types d'informations. Il n'y a pas de stockage dans l'administrateur central. La boucle phonologique – *phonological loop* – est l'un d'eux : ce système stocke et classe les informations de nature phonologique, ou plus largement de nature langagière. Elle est aussi responsable des processus internes d'autorépétition (*rehearsal process*). Cette composante de la MdT a été mise en relation avec de nombreux phénomènes psycholinguistiques (Delafoy et Ehrlich, 1990 ; Fortier, 2013). Le calepin visuo-spatial – *visuo-spatial sketchpad* – est physiologiquement distinct de la boucle phonologique. Il est responsable de l'emmagasinage d'informations visuo-spatiales et de la manipulation d'images mentales (Baddeley, 1986 ; Delafoy et Ehrlich, 1990). Le buffer épisodique (*episodic buffer*) complète l'architecture du modèle de Baddeley. Cette dernière composante a été rajoutée au modèle initial de la MdT afin de conceptualiser l'accès en MLT à du matériel accessible à la MdT, et vice versa (Baddeley, 2000). Elle constitue un système de stockage limité en temps et en éléments, au même titre que les autres sous-systèmes esclaves (Alamargot et Chanquoy, 2002). Cependant, le buffer épisodique manipule des informations de natures diverses – sémantiques, conceptuelles, visuo-spatiales, phonologiques, etc. – et permet l'échange d'informations entre les deux sous-systèmes et la MLT. Il a la particularité de stocker des « épisodes », entendus comme des éléments d'informations spatiales et temporelles (Alamargot et Chanquoy, 2002).

Il s'ensuit que les composantes langagières du modèle sont la boucle phonologique et l'administrateur central. Selon la nature de l'épreuve de mémoire, ce sont l'un, l'autre ou les deux ensembles qui sont mesurés. Par exemple, plusieurs auteurs suggèrent que la répétition de non-mots constitue une mesure de la boucle phonologique, donc un sous-ensemble de la MdT (French, Simard et Fortier, 2012 ; Poncelet et Van der Linden, 2003). Gathercole et Baddeley (1996) ont proposé que cette épreuve constitue une mesure relativement pure de ce qu'ils appelaient la mémoire à court terme verbale (*verbal short-term memory*). Cependant, si les non-mots contiennent des syllabes qui sont en elles-mêmes des unités lexicales, ou des segments qui sont possibles et fréquents en français, les connaissances lexicales sont réputées avoir un effet sur le rappel. Toujours est-il que la répétition de non-mots requiert du sujet qu'il n'utilise que sa capacité à stocker temporairement et rappeler des représentations phonologiques pour réussir l'épreuve, étant donné que les connaissances lexicales stockées en MLT ne sont pratiquement d'aucun renfort (Archibald et Gathercole, 2007).

Ainsi, la boucle phonologique, le calepin visuo-spatial et, dans une certaine mesure, le buffer épisodique, reprennent les composantes de la mémoire à court terme, telle que décrite dans les travaux antérieurs (Engle *et al.*, 1999). Baddeley fait ainsi fonctionner les composantes de la MCT (stockage et rappel inchangé d'éléments) à l'intérieur de son modèle de la mémoire de travail. Ainsi, les articles post-Baddeley (1986) évoquent plutôt l'idée de « mémoire à court terme verbale », donc de la boucle phonologique, lors de l'évaluation de l'habileté à stocker puis à reproduire des séquences de phonèmes dans l'ordre où ils ont été présentés (Martinez Perez *et al.*, 2012).

2.2.2 Retour sur la mémoire à court terme

Les épreuves que nous allons utiliser dans la partie empirique, soit la répétition de mots rares, de chiffres, de phrases et de rythmes, fonctionnent sur un rappel sériel et sans traitement d'éléments de différentes natures, ce qui les positionne du côté de la MCT. Nous devons ainsi nous extraire de la MdT de Baddeley et retourner aux études de la MCT de Atkinson et Shiffrin (1968), de Atkinson et Shiffrin (1971) et de Peterson et Peterson (1959) pour préciser ce cadre théorique. La MCT constitue le système qui se charge du stockage passif et du rappel de petites quantités d'informations, comme des mots ou des chiffres. L'article princeps de Miller (1956) pose les prémisses de la MCT et suggère que le rappel d'éléments devient plus problématique au-delà du chiffre « magique » de 7 éléments, plus ou moins 2. Il introduit également, dans le même article, la notion de tronçons d'informations (*chunks*). Cette notion permet d'expliquer qu'un sujet parvienne à retenir des informations regroupées. Cela réduit le nombre d'éléments à retenir et réduit la charge en MCT. Ainsi, 7 éléments sont retenus généralement par un adulte, dans une tâche classique d'empan de mémoire.

Par opposition, Baddeley (2000) propose que la limite est surtout temporelle, plutôt que liée au nombre d'items. Le délai entre le stimulus et la répétition doit être court. La MCT est d'une durée limitée : l'information disparaît si elle n'est pas utilisée (Peterson et Peterson, 1959). L'« empan de mémoire » correspond à cet intervalle de temps entre le stimulus et le rappel où l'information est maintenue passivement et où le rappel est possible parce que les traces en mémoire sont toujours disponibles. Dowling (1973) a observé que le rappel d'éléments rythmiques présentait les mêmes patrons que le rappel d'éléments linguistiques (*chunking*, effets de primauté et de récence) et se rapportait donc à la MCT. Dans la vie quotidienne, la MCT est censée être mise à contribution lorsque l'on tente de répéter un numéro de téléphone tout juste après l'avoir entendu (Atkinson et Shiffrin, 1971). Les tâches typiquement utilisées pour évaluer la MCT sont la répétition de mots et de chiffres (Daneman et Carpenter,

1980). Le tableau 2.1 permet de bien délimiter la MCT, en la contrastant avec les caractéristiques associées à la MdT.

Tableau 2.1 – Comparaison des caractéristiques de la MdT et de la MCT

| Caractéristiques | Mémoire de travail (MdT) | Mémoire à court terme (MCT) |
|-----------------------------|--|--|
| Nature des processus | Stockage et traitement de l'information, processus attentionnels. | Stockage passif de l'information, empan de mémoire : 7 +1 2 éléments, sensible aux effets de primauté et de récence. |
| Nature des tâches associées | Effectuer une tâche cognitive tout en maintenant des informations emmagasinées | Effectuer un rappel des éléments produits, dans l'ordre dans lequel ils ont été présentés |
| Épreuve classique | Empan de lecture (Daneman et Carpenter, 1980) | Répétition de mots, répétition de chiffres. |
| Exemple dans le quotidien | Maintien en mémoire d'une adresse tout en écoutant les instructions pour s'y rendre. | Maintien en mémoire d'un numéro de téléphone. |

Dans le cadre de ce mémoire de maîtrise, le libellé MCT est appliqué à l'habileté qui est mise à contribution lors d'épreuves classiques de répétition de chiffres et de mots, dans la mesure où elle est ainsi abordée dans les études plus récentes (Martinez Perez *et al.*, 2012 ; Melby-Lervåg, 2011 ; Parrila *et al.*, 2004 ; Swanson et Berninger, 1996).

2.2.3 Les mémoires en lien avec la lecture

Plusieurs modèles portant sur la compréhension en lecture font appel à la notion de capacité (Daneman et Carpenter, 1980 ; Just et Carpenter, 1992). Cette idée se réfère au fait que le traitement d'informations est de nature limitée, et que ces limites sont

source de variation interindividuelle. Ainsi, la mémoire de travail (Daneman et Carpenter, 1980) et la mémoire à court terme ont toutes deux été corrélées avec les performances en lecture (compréhension) et en décodage de mots (Swanson et Siegel, 2001). Plusieurs travaux récents attribuent les différences observées entre les bons et les mauvais lecteurs à la MdT plutôt qu'à la MCT (Gathercole, 2009 ; Just et Carpenter, 1992). En effet, les résultats de Cain *et al.* (2004a) et d'autres études plus récentes suggèrent que c'est effectivement la MdT qui constitue l'entité où certains processus liés à la compréhension, comme les inférences, ont lieu. Cette assertion a été maintes fois vérifiée, tant chez les enfants (Cain *et al.*, 2004a ; Cain *et al.*, 2004b ; Seigneuric *et al.*, 2000 ; Swanson et Berninger, 1995 ; Swanson et Howell Ashbaker, 2000) que chez les adultes (Daneman et Carpenter, 1980 ; Engle, Cantor et Carullo, 1992 ; Engle *et al.*, 1999 ; Just et Carpenter, 1992). Rappelons que la MdT, conçue comme l'habileté permettant la coordination de deux types de processus, soit le stockage et le traitement, implique des processus plus dynamiques que la MCT, qui n'est responsable que du stockage et dont les processus en jeu sont plutôt statiques. Il existe néanmoins des preuves empiriques selon lesquelles la MCT est corrélée avec les capacités à répéter des phrases, mais moins avec la compréhension de ces phrases. Le patron général est que les mauvais lecteurs semblent réussir moins bien que les bons lecteurs aux tâches qui demandent un stockage à court terme d'informations verbales (Swanson et Howell Ashbaker, 2000).

Dans les dernières années, c'est surtout la boucle phonologique, sous-système de stockage et de rappel d'items phonologiques dans la MdT, qui a attiré l'attention des chercheurs. Certains auteurs ont suggéré que la boucle phonologique faisait référence à la mémoire à court terme (Baddeley, 1986 ; Fortier, 2013). Baddeley (1986) en appelle à la similarité de leurs composantes : stockage d'information verbale et processus internes de répétition. La capacité à répéter des non-mots, une mesure classique de la boucle phonologique, a été corrélée avec les mesures de compréhension en lecture (Kibby, 2009 ; Poncelet et Van der Linden, 2003 ; Swanson et Siegel,

2001). De la même façon, ce que certains auteurs appellent la MCT verbale a pris une importance prépondérante dans la littérature sur les préalables cognitifs à l'acquisition de la lecture (Lieberman et Shankweiler, 1989 ; Martinez Perez *et al.*, 2012). L'observation de patients présentant un trouble du langage et de la compréhension et de faibles scores dans les tâches faisant appel à la MCT est à la source de l'intuition de cette relation (Gathercole, 2009). L'interprétation de ces résultats ne fait toutefois pas consensus : d'un côté, les habiletés en MCT sont impliquées dans les performances en lecture ; de l'autre, la MCT constitue plutôt un sous-ensemble d'un processus plus général de traitement phonologique, qui influence la lecture. Lieberman et Shankweiler (1989) considèrent pour leur part qu'une faible MCT est susceptible d'influencer la compréhension, dans la mesure où elle est impliquée dans le traitement du texte – les auteurs insistent sur son rôle dans la segmentation phonologique. Selon eux, les mauvais lecteurs sont à la merci d'une MCT plus facilement surchargée lors de la lecture de textes, ce qui constitue un frein principal à leur compréhension, rappelant les conclusions de Just et Carpenter (1992). Perfetti et Lesgold (1977), qui souscrivaient à une modélisation binaire de la mémoire (MLT et MCT), ont imputé à la MCT un rôle important dans les processus de compréhension en lecture. Les auteurs proposent que la MCT participe aux segmentations phonologique et lexicale ainsi qu'aux processus de lecture de bas niveau, influençant ainsi les performances en compréhension : « *It is clear that STM can become a bottleneck for higher-level processing solely because of the speed of coding into and out of STM* » (Perfetti et Lesgold, 1977, p. 59). L'influence de cette habileté plus statique sur des processus de lecture de bas et de haut niveau est à vérifier.

2.3 Le vocabulaire

2.3.1 Le vocabulaire et les propriétés des mots

Le lexique mental se réfère à l'entité théorique faisant partie de la MLT qui regroupe l'ensemble des mots qu'un individu connaît ainsi que leurs propriétés phonétiques, graphiques, phonologiques, syntaxiques et sémantiques (Elman, 2004 ; Levelt, 1989 ; Taft, 1992). Il constitue un ensemble de réseaux de mots, très organisé, qui permet aux lecteurs d'accéder à de la signification lorsqu'un input visuel ou auditif est perçu (Coltheart, 2004). Ainsi, les mots connus par un individu et leurs propriétés sont réputés faire partie de la MLT. On peut soutenir qu'un mot connu est un mot dont les formes orale et graphique, le sens (usage, concepts connexes) et l'usage (registre, fréquence) sont contenus en MLT (Qian, 2002). Bogaards (2000) préfère l'appellation « unité lexicale », et dresse une liste des caractéristiques qui doivent être rencontrées afin de connaître un mot : forme, signification, morphologie, syntaxe, collocation et discours. La plupart des définitions de la connaissance lexicale se réfèrent à la disponibilité en mémoire du mot et de ses propriétés intrinsèques. Une exposition récurrente à un nouveau mot est requise pour en apprendre les propriétés (Morris et Labelle, 2008). La structure de l'entrée lexicale de Levelt (1989), reproduite dans la figure 2.3, illustre bien cette idée de propriétés des mots.

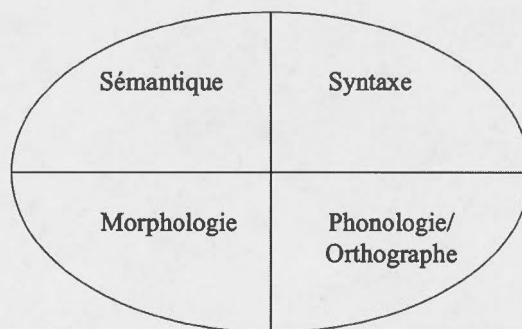


Figure 2.3 – La structure de l'entrée lexicale – adapté de Levelt (1989) (cité dans Jiang, 2000)

L'idée de représentation lexicale implique que les informations relatives à chaque entrée sont immédiatement accessibles et activées simultanément. Comme le relate

Jiang (2000), les différentes représentations en compétition pour un mot ambigu sont toutes activées, même dans un contexte induisant une interprétation plutôt qu'une autre. La complexité du processus entourant le tissage de ces liens entre les propriétés des entrées lexicales et entre les entrées lexicales elles-mêmes se développe grâce à une exposition langagière riche. L'intégration de l'ensemble de ces propriétés pour chaque entrée et leur activation simultanée garantissent une utilisation efficace du langage, tant en production qu'en perception (Jiang, 2000).

Deux notions ont pris de l'importance dans la littérature sur le vocabulaire ces dernières années : les termes « profondeur » et « étendue » sont maintenant utilisés pour évaluer les connaissances lexicales (Qian, 1999 ; Read, 1988). Ces deux dimensions de la maîtrise du vocabulaire se sont avérées d'égale importance dans la prédiction des performances en lecture académique (*academic reading*) et en compréhension de la lecture (Qian, 2002). L'étendue du vocabulaire renvoie généralement au nombre de mots connus, au moins de façon superficielle, par un individu, que Miller (1996) a estimé à 45 000 chez un adulte. La profondeur, quant à elle, est la dimension qui tente d'évaluer la qualité du vocabulaire, c'est-à-dire la connaissance de certaines propriétés des mots – phonologie, phonétique, sémantique, collocations, registre, fréquence, par exemple – et des liens entre celles-ci (Qian, 2002). Cette conception multidimensionnelle du vocabulaire a pour avantages de mieux circonscrire le sujet d'étude et de donner l'occasion aux chercheurs de développer de meilleurs outils de mesure du vocabulaire et de ses éléments constitutifs.

Le construit théorique qu'est le lexique mental, très utilisé en psycholinguistique, ne fait cependant pas consensus dans la littérature scientifique (Rastle, 2009). L'étendue et la profondeur du vocabulaire participe néanmoins massivement de la compétence langagière générale des sujets (Alloway et Gathercole, 2005). Par exemple, la

littérature sur les langues secondes considère le vocabulaire comme un sous-concept faisant partie des habiletés langagières réceptives et expressives chez un individu, ou *oral language proficiency* (August et Shanahan, 2006).

2.3.2. Le vocabulaire en lien avec la lecture

Le lien de corrélation entre le vocabulaire et la compréhension en lecture est bien établi, tant chez les enfants que chez les adultes (Baumann, 2005 ; Cain *et al.*, 2004b ; Freebody et Anderson, 1983 ; Lervåg et Aukrust, 2010 ; Qian, 1999). La littérature sur le vocabulaire et la compréhension a eu recours à des modèles psycholinguistiques pour expliquer les corrélations et les prédictions observées. Ainsi, le lien qui lie le vocabulaire à la compréhension est indéniable et attesté, mais demeure mal expliqué, en partie à cause de la complexité inhérente à ces deux processus. De plus, la direction du lien de causalité ne fait pas consensus. Nous entendons par causalité la contribution théorique d'une habileté à une mesure finale, et non pas un simple lien de corrélation. Les positions à ce sujet se divisent en trois groupes : soit le vocabulaire influence la performance en lecture (Hemphill et Tivnan, 2008 ; Lervåg et Aukrust, 2010 ; Qian, 2004), soit le fait de lire contribue positivement au vocabulaire (Cain et Oakhill, 2011) ou soit il existe un même processus à la source du développement de ces deux habiletés. Nagy (2005) a notamment proposé un lien de causalité bidirectionnel, où chacune des deux composantes est nourrie par l'autre.

La première position, qui propose que le vocabulaire permet la compréhension en lecture par l'entremise d'un lien direct et causal, correspond à ce que nous souhaitons mesurer dans la présente étude et se défend bien. Il est certainement raisonnable de penser que la richesse du vocabulaire et l'accès rapide aux mots et à leurs propriétés a une influence positive sur la compréhension de phrases et de texte. Autrement dit, les mots d'un texte doivent être connus pour que la compréhension ne soit pas

hypothéquée par une méconnaissance du vocabulaire (Lervåg et Aukrust, 2010 ; Nation, 2001). Cette position forte, défendue notamment par Anderson et Freebody (1981), bénéficie de preuves empiriques (Stahl, 1999). Elle a évidemment ses détracteurs : Freebody et Anderson (1983) (cité dans Cain *et al.*, 2004b), notamment, ont observé qu'un vocabulaire limité n'était pas obligatoirement corrélé avec des déficits en compréhension.

Si des travaux en psychologie, en linguistique et en sciences cognitives font état de la corrélation qui existe entre les habiletés en vocabulaire et les performances en compréhension de la lecture en L1 (Carroll, 1993 ; Hemphill et Tivnan, 2008 ; Lervåg et Aukrust, 2010 ; Ricketts, Nation et Bishop, 2007 ; Shiotsu et Weir, 2007 ; Tunmer et Chapman, 2012), des études en langues secondes ont montré que le vocabulaire était un prédicteur plus important de la compréhension chez les apprenants L2 que chez les enfants qui lisent dans leur langue maternelle (Lervåg et Aukrust, 2010). Ces auteurs ont en effet conclu que les habiletés en vocabulaire chez les locuteurs L2 expliquent une grande part de la variance entre les locuteurs L1 et L2 pour ce qui est des performances en compréhension de la lecture et ce, tant initialement que plus tard dans le développement. De la même façon, les enfants qui sont faibles en compréhension de la lecture présentent des faiblesses lorsque vient le temps d'inférer le sens de mots nouveaux à partir du matériau textuel (Cain *et al.*, 2004b). Le fait de posséder un vocabulaire étendu avant même de fréquenter l'école peut influencer positivement le développement langagier ultérieur : cela sous-tend un vocabulaire plus riche plus tard, et mène à un taux plus élevé du développement de la compréhension en lecture (Anderson et Freebody, 1981). L'étude longitudinale de Hemphill et Tivnan (2008) est également sans équivoque : le vocabulaire est un prédicteur constant des habiletés en lecture, tandis que les qualités prédictives du décodage et de la conscience phonologique subissent un déclin avec le temps (Spira, Bracken et Fischel, 2005). Cette position instrumentaliste de la relation vocabulaire-compréhension est d'ailleurs mise de l'avant dans plusieurs écoles, où on procède à un

enseignement de mots de vocabulaire avant de passer à une activité de compréhension d'un texte contenant les mots appris. Ce procédé didactique a prouvé son efficacité (Beck et McKeown, 1991 ; cité dans Cain *et al.*, 2004b).

Tandis que la relation entre le vocabulaire et la compréhension est attestée, celle qui lie le vocabulaire au décodage, ce dernier conçu en tant que mesure finale et prédite, n'a pas particulièrement retenu l'attention des chercheurs. Il est effectivement moins certain que les habiletés en vocabulaire jouent un rôle dans les processus de lecture de plus bas niveau, comme le décodage. L'étude de Ricketts *et al.* (2007) propose de remédier à cette situation. Ils font l'hypothèse que la connaissance du vocabulaire sur un plan sémantique permet le développement de représentations grapho-phonologiques plus fines, conduisant par exemple un enfant à reconnaître plus rapidement des mots dont l'orthographe est plus opaque (Plaut, McClelland, Seidenberg et Patterson, 1996 ; cité dans Ricketts *et al.*, 2007). L'idée derrière cette affirmation est que la lecture de mots opaques, dont l'orthographe n'est pas consistante avec la suite phonologique, doit se faire autrement que par le décodage uniquement : elle doit faire intervenir une composante sémantique. Ainsi, un enfant présentant des bonnes habiletés en vocabulaire serait plus enclin à affiner et à mieux organiser les représentations phonologiques et graphiques des entrées lexicales connues, rendant efficace la reconnaissance visuelle de mots. Si Ricketts *et al.* (2007) ont bien trouvé que le vocabulaire contribuait à la variance en lecture de mots avec des inconsistances grapho-phonologiques, ils ont finalement échoué à prouver que le vocabulaire pouvait expliquer en partie le décodage.

Ces assertions font écho au modèle DRC de Coltheart *et al.* (2001). En effet, le modèle de reconnaissance visuelle de mots inclut une composante sémantique : la route lexicale. Cependant, la voie indirecte du modèle, qui fonctionne grâce à un système de règles de conversions graphèmes-phonèmes, fait directement référence au décodage et ne fait pas intervenir de composante lexicale ou sémantique. Cela n'est pas compatible

avec l'idée que la variance en décodage peut être en partie tributaire des habiletés en vocabulaire : le décodage est hors-vocabulaire par définition.

2.4 Synthèse

La recension des écrits nous a permis de présenter les construits théoriques MCT et vocabulaire et de les mettre en relation avec la compréhension en lecture et le décodage. Les parties qui suivent permettront de mesurer et modéliser l'influence, chez les jeunes Innus de Pessamit, des deux habiletés, d'ordres cognitif et langagier, sur la lecture.

CHAPITRE III

MÉTHODE

Le présent travail de recherche vise d'abord à déterminer quelles épreuves, parmi celles ayant été soumises aux participants à trois moments différents, diagnostiquent le mieux les problèmes ultérieurs en lecture, pour ensuite explorer la contribution relative des habiletés qui se dégagent des épreuves aux performances en lecture. Ce chapitre présente la méthode utilisée pour répondre à ces deux questions de recherche. Nous commencerons par présenter notre corpus, le fonctionnement des différentes collectes de données et les participants (3.1). Par la suite, nous décrirons les instruments de mesure pour les temps 1, 2 et 3 (3.2). Les épreuves qui nous ont permis de constituer nos données au temps final seront décrites dans la section 3.3. Nous procéderons à une synthétisation des variables et des épreuves dans la section suivante (3.4). Nous conclurons avec une brève description des traitements statistiques utilisés pour répondre à nos questions de recherche (3.5).

3.1 Description du corpus

Notre recherche, de type longitudinal, vise, dans un premier temps, à vérifier la valeur prédictive de certaines épreuves aux performances ultérieures en lecture. Cette question est envisagée dans une optique diagnostique. Les variables indépendantes correspondent donc aux résultats aux épreuves, mesurées à trois périodes différentes, et les variables dépendantes sont les performances en lecture, en décodage et en compréhension au temps final. Dans un deuxième temps, nous cherchons à évaluer la

contribution des habiletés en mémoire à court terme et en vocabulaire, telles que mesurées en trois temps, aux performances en lecture au temps final. Dans cette équation, les habiletés en mémoire à court terme et en vocabulaire correspondent aux variables indépendantes, et les performances en lecture à la variable dépendante. Ce deuxième objectif est de nature explicative.

3.1.1 Collectes de données

Notre recherche s'inscrit dans le cadre d'une étude plus large qui vise à dresser un portrait de la littératie et des habiletés langagières en innu et en français dans différentes communautés autochtones, dont celle de Pessamit. Lori Morris, professeure au Département de linguistique de l'Université du Québec à Montréal, est la chercheuse principale de ce projet. Le corpus est constitué de données recueillies en quatre temps, sur une période de cinq années au total, chez 76 enfants fréquentant l'école de conseil de bande de la communauté autochtone de Pessamit. La communauté est située entre Forestville et Baie-Comeau, sur la Côte-Nord, au Québec. Les enfants avaient 5 ans lors de leur première participation aux épreuves. Ils appartiennent à deux cohortes : la cohorte 2008-2009 et la cohorte 2009-2010. Ces élèves ont participé aux épreuves chaque année à partir de la maternelle, soit en 2008 pour la cohorte 2008-2009 et en 2009 pour la cohorte 2009-2010. Toutes les cohortes ont été évaluées au début et à la fin de la maternelle et une fois par année (au printemps) par la suite. La collecte de données pour évaluer les capacités en lecture des deux cohortes a eu lieu simultanément, au printemps 2013.

Les données des deux cohortes sont regroupées selon l'âge et le moment dans la vie scolaire des élèves, au moyen de quatre libellés qui correspondent à quatre périodes de temps. Au temps 1 (désormais T1), tous les enfants sont au début de la maternelle et ont 5 ans (moyenne : 5,55 ans; écart-type : 0,28 an; n = 50). Au temps 2 (désormais

T2), les enfants sont à la fin de leur année scolaire en maternelle et ont environ 6 ans (moyenne : 6,09 ans; écart-type : 0,27 an; n = 48). Au temps 3 (désormais T3), les enfants ont 7 ans (moyenne : 7,10 ans; écart-type : 0,26 an; n = 53) et sont tout près de terminer leur première année du primaire. Au temps final (désormais TF), qui correspond à la collecte de données effectuée au printemps 2013 auprès des deux cohortes simultanément, les enfants en sont alors à leur 3^e ou 4^e année du primaire et ont entre 9 et 10 ans (moyenne : 9,53 ans; écart-type : 0,52 an; n = 70). Les données des temps 1, 2 et 3 ont été recueillies par M^me Lori Morris et son équipe. L'auteure de ce mémoire a procédé à la cueillette des données du temps final.

3.1.2 Participants et exclusions

Il est à noter que tous les enfants n'ont pas participé aux collectes de données effectuées à tous les temps. En effet, 46 enfants ont été soumis aux épreuves à chacun des quatre temps (T1, T2, T3 et TF). Nous possédons certaines données des autres enfants, mais pas toutes, pour cause de maladie, de départ de la communauté ou d'autres raisons. Les participants ne sont donc pas exactement tous les mêmes à tous les temps, ce qui ne constitue en rien un obstacle pour effectuer les statistiques de haut niveau auxquelles nous souhaitons procéder.

Les élèves présentant de graves problèmes intellectuels ou sensori-moteurs ont été exclus d'office de l'étude. Nous avons remarqué que les participants 25 et 13 présentaient des résultats aberrants sur un total de quatre épreuves chacun. En effet, les résultats de ces élèves à ces épreuves s'écartaient de la moyenne par deux écarts-types. Par ailleurs, les résultats de ces participants étaient systématiquement bas dans le reste des épreuves. De plus, sur l'ensemble des quatre temps, nous ne pouvions observer aucune amélioration. Étant donné qu'il était raisonnable de soupçonner un

important retard dans le développement cognitif de ces deux participants, nous avons décidé de les exclure de l'analyse.

3.2 Description des épreuves des temps 1, 2 et 3

Les épreuves utilisées au temps 1, 2 et 3 proviennent toutes des Nouvelles Épreuves pour l'examen du langage (désormais NÉEL) (Chevrie-Muller et Plaza, 2001), une batterie d'évaluation du langage oral des enfants de 3 ans 7 mois à 8 ans 7 mois conçue pour des orthophonistes (Chevrie-Muller et Plaza, 2001). La batterie est constituée de sous-tests qui étudient diverses composantes du langage. Elle se divise en deux formes : la forme P est réservée aux plus jeunes enfants, soit ceux de 3 ans 7 mois à 6 ans 6 mois, tandis que la forme G vise un public plus âgé, soit les enfants de 6 ans 6 mois à 8 ans 7 mois). Dans le projet de recherche, les épreuves de la forme P ont été utilisées aux temps 1 et 2, tandis qu'au temps 3, ce sont les épreuves de la forme G qui ont été soumises aux élèves. Certaines épreuves, ou sous-tests, sont communs à l'une et l'autre forme, mais se présentent sous une version augmentée dans la forme G. Par exemple, l'épreuve de répétition de chiffres au temps 3 (forme G du NÉEL) comporte une série de chiffres de plus que l'épreuve du même nom utilisée aux temps 1 et 2 (forme P du NÉEL).

Le temps alloué à la passation de l'ensemble des épreuves était d'environ 25 à 30 minutes par élève. Les épreuves ont toujours été soumises dans l'ordre suivant : répétition de mots rares, répétition de chiffres, répétition de rythmes (au temps 3 exclusivement) topologie et arithmétique, vocabulaire (expression), répétition de phrases et vocabulaire (réception). Les passations, individuelles, ont généralement eu lieu dans le corridor à l'extérieur de la classe.

Les instruments de mesure de l'ensemble de la batterie NÉEL ont été testés sur une population de 541 élèves français âgés de 3 ans 7 mois à 8 ans 7 mois fréquentant des écoles maternelles et élémentaires dans les années 1998-1999 et 1999-2000. Une comparaison statistique des résultats obtenus par ces enfants par rapport à ceux obtenus dans la communauté de Pessamit figure d'ailleurs à l'Annexe B.

Les expérimentatrices qui se sont chargées de la passation des épreuves dans la communauté étudiée n'ont fait passer que certains sous-tests aux enfants. Afin de mener à bien la présente étude, nous avons sélectionné les sous-tests qui permettaient de mesurer le vocabulaire et la mémoire à court terme. Les instruments de mesure pour les temps 1, 2 et 3, décrits dans la prochaine section de notre mémoire, sont donc tous tirés du NÉEL. Tous les pointages sont aussi issus des consignes du NÉEL. Nous nous efforcerons de décrire les habiletés que ces épreuves sont réputées mobiliser.

3.2.1 Épreuves mesurant le vocabulaire

Nous présenterons dans cette section les épreuves Topologie et arithmétique, Vocabulaire expressif et Vocabulaire réceptif.

3.2.1.1 Topologie et arithmétique

Les sous-tests suivants ont été regroupés sous la bannière « Topologie et arithmétique » dans notre étude. Il est à noter que les sous-tests du temps 3 diffèrent de ceux des temps 1 et 2.

Topologie et arithmétique. Cette épreuve de compréhension ne nécessite pas de réponse orale de la part du participant. Celui-ci doit manipuler une figurine de chat

par rapport à une figurine de chaise selon les instructions fournies par l'expérimentatrice. L'élève est ainsi amené à prouver sa compréhension de 7 prépositions, comme les termes « à côté de » ou « devant ». L'épreuve requiert aussi que l'enfant s'oriente dans l'environnement qui lui est imposé et visualise ce qu'il doit faire (Chevrie-Muller et Plaza, 2001, p. 92). Un total de 12 points est alloué pour cette épreuve. Chaque bonne réponse vaut 2 points, sauf le dernier item, « en dessous de », où 1 point est accordé en cas d'échec à l'item « sous ». Ce sous-test est soumis aux participants aux temps 1 et 2.

Arithmétique 1. Cette épreuve vise à évaluer la maîtrise des notions de rang et de nombre (Chevrie-Muller et Plaza, 2001, p. 92). L'enfant doit prouver sa compréhension des consignes sous trois modalités de réponses. D'abord, le participant est amené à pointer du doigt la figurine demandée parmi celles disposées en rang devant lui – « Montre-moi le premier », par exemple. On lui demande ensuite de dire combien d'animaux, puis combien de chats on retrouve dans le groupe de figurines. L'élève est finalement sommé de mettre « les animaux deux par deux ». Cette épreuve mesure la compréhension de certaines opérations mathématiques de base (Chevrie-Muller et Plaza, 2001, p. 92). On attribue 2 points à chaque item correctement identifié ou 0, le cas échéant, pour un total sur 10. Ce sous-test est soumis aux participants aux temps 1 et 2.

Arithmétique 2. Aucune réponse orale n'est nécessaire pour la réussir l'épreuve, qui vise à évaluer la maîtrise des notions de quantité. Quatre boîtes, contenant respectivement 3, 9, 3 et 2 jetons, sont présentées à l'enfant. Ce dernier doit démontrer sa compréhension de certains items lexicaux, en montrant du doigt la boîte sous-entendue par la consigne (« Montre-moi la boîte où il y a le moins de jetons », par exemple). Pour chaque item correctement identifié, l'enfant obtient un point, pour un total de 4. Ce sous-test est soumis aux participants aux temps 1 et 2.

Topologie 2. Cette épreuve vise à évaluer la maîtrise de certaines notions liées à l'espace. Le participant doit manipuler des figurines de canards, en les mettant « côte-à-côte » ou « l'un derrière l'autre », par exemple. On attribue 2 points à chacune des 7 manipulations correctement effectuées ou 0, le cas échéant, pour un total sur 14. Ce sous-test est soumis aux participants au temps 3 seulement.

Topologie 3. Cette épreuve vise à évaluer la maîtrise de certaines prépositions et des consignes liées à l'espace. Le participant doit manipuler deux figurines de chats et une figurine de canard (ex. « Mets un chat au milieu du jardin »). Il doit aussi identifier du doigt l'animal selon la consigne (ex. « Montre-moi l'animal qui est le plus loin du jardin »). On attribue 2 points à chacune des 8 manipulations et identifications effectuées correctement ou 0, le cas échéant, pour un total sur 16. Ce sous-test est soumis aux participants au temps 3 seulement.

3.2.1.2 Vocabulaire (expression)

Les deux épreuves qui construisent la variable Vocabulaire (expression) dans notre recherche se caractérisent par les mêmes modalités de réponse. Dans les deux cas, le participant doit prononcer le mot qui correspond à l'image lui étant présentée. Chevrie-Muller et Plaza (2001, p. 104) considèrent que les deux épreuves d'expression du vocabulaire permettent de dresser un portrait du « stock lexical et [de l']accès à ce stock ». Les épreuves sont utilisées au temps 1 et 2 et figurent sous une version augmentée au temps 3.

Dénomination – vocabulaire 1. Cette épreuve cherche à mesurer l'étendue lexicale de l'enfant (Chevrie-Muller et Plaza, 2001). On présente à l'enfant une série d'images, soit 14 aux temps 1 et 2 et 36 au temps 3, représentant chacune un mot « concret » à identifier. Deux points sont alloués pour chaque bonne réponse, pour un total de 28 aux temps 1 et 2 et de 72 au temps 3. Lorsque l'enfant est incapable d'identifier

l'image, le premier phonème du mot recherché est fourni à l'enfant par l'expérimentatrice. Une bonne réponse à la suite de cette ébauche orale vaut 1 point. Cette tâche évalue un ensemble de capacités, comme la présence du mot dans le lexique mental, l'accès au lexique ou la capacité à produire le mot, toujours selon Chevrie-Muller et Plaza (2001).

Dénomination – vocabulaire 2. Cette épreuve permet d'évaluer, encore une fois, l'étendue lexicale de l'élève. Dans l'épreuve *Dénomination – vocabulaire 2*, cependant, les images à identifier correspondent à des couleurs, des formes et des parties du corps. Les séries des couleurs et des formes comportent chacune 6 images à identifier, comme pour l'épreuve *Dénomination – vocabulaire 1*. Les 8 parties du corps à nommer sont quant à elles désignées par l'expérimentatrice sur son propre corps. Les consignes à propos de l'ébauche orale restent identiques à celles qui prévalent dans l'épreuve *Dénomination – vocabulaire 1*. L'épreuve originale comprenait 21 items à identifier. Dans les données dont nous disposons, l'item « index » a été systématiquement exclu des résultats, puisqu'il n'a pas été évalué de façon constante par les différentes expérimentatrices, certaines acceptant la réponse « doigt », et d'autres non. Ceci porte à 20 le nombre d'items à identifier à tous les temps, pour un total de 40 points.

3.2.1.3 Vocabulaire (réception)

Les deux épreuves qui nous permettent de calculer notre variable Vocabulaire (réception) fonctionnent selon une même modalité de réponse. En effet, dans ces épreuves, le participant a pour consigne de pointer l'image ou la partie du corps qui représente le mot prononcé par l'expérimentatrice. Dans la mesure où nous nous intéressons à l'étendue du vocabulaire sur le plan réceptif et où les deux épreuves

ciblent simplement des thématiques différentes, nous avons comptabilisé les deux épreuves ensemble.

Il est à noter que les 14 items qui étaient à identifier dans l'épreuve *Dénomination – vocabulaire 1* sont également à identifier dans l'épreuve *Réception – vocabulaire 1*, où 22 autres items s'ajoutent. Les items de l'épreuve *Dénomination – vocabulaire 2* sont, quant à eux, repris intégralement dans l'épreuve *Réception – vocabulaire 2*. Rappelons que le procédé de passation de la batterie prévoit que les épreuves d'expression précèdent les épreuves de réception, et que jamais l'élève n'est corrigé pendant l'expérimentation. Ainsi, l'écueil du non-respect du principe de l'indépendance des observations est évité.

Réception – vocabulaire 1. Cette épreuve cherche à mesurer l'étendue lexicale de l'élève. On lui présente une planche comprenant 8 images. L'expérimentatrice dit un mot qui correspond à l'une des images, et l'enfant doit l'identifier du doigt. Cette tâche évalue un ensemble de capacités cognitives, comme le décodage visuel et la présence du mot dans le lexique mental, selon Chevrie-Muller et Plaza (2001). L'épreuve est composée de 36 items et notée sur 36 points, et ce, aux temps 1, 2 et 3.

Réception – vocabulaire 2. L'épreuve permet de mesurer les connaissances lexicales de l'enfant sur le plan de la réception. Ce dernier est exposé à une série de planches comportant 8 images et doit identifier celle qui correspond la couleur ou la forme nommée par l'expérimentatrice. Les 8 parties du corps à pointer sont quant à elles désignées par l'élève sur son propre corps. L'item « index » a été encore une fois exclus des résultats finaux. L'épreuve est donc notée sur 20 points pour les temps 1, 2 et 3.

3.2.2 Épreuves mesurant la mémoire à court terme

Les habiletés mnésiques font l'objet d'une évaluation à l'aide de différentes épreuves impliquant la répétition ordonnée d'éléments, soit les épreuves de répétition de chiffres, de mots, de phrases (T1 et T2) et de rythmes (T3).

3.2.2.1 Répétition de mots rares

La tâche de répétition de mots rares requiert que l'enfant perçoive le mot, le stocke temporairement en mémoire puis reproduise la séquence entendue (Chevrie-Muller et Plaza, 2001, p. 60). Les auteurs du NÉEL affirment que les mots choisis ne font pas partie du lexique mental des élèves, étant donné leur rareté. Ils attribuent aux mots polysyllabiques choisis les propriétés de non-mots (Chevrie-Muller et Plaza, 2001, p. 61). L'épreuve de répétition de non-mots est une tâche classique d'évaluation de la boucle phonologique dans la mémoire de travail (Parrila *et al.*, 2004 ; Poncelet et Van der Linden, 2003). L'enfant doit répéter trois séries de 3 mots allant de 3 à 5 syllabes aux temps 1 et 2. Au temps 3, une série s'ajoute, comportant des mots de 6 syllabes. Puis, un deuxième ensemble de mots doit être répété sous le même protocole. Ce deuxième ensemble introduit des difficultés phonologiques, telle une « succession de consonnes à la frontière syllabique » (Chevrie-Muller et Plaza, 2001, p. 61), comme /g/ et /z/ dans le mot *exégèse*, ou encore des clusters phonologiques, par exemple la syllabe /struk/, dans l'item *obstruction*. La note attribuée est de 1 point par item si l'enfant répète le mot entièrement et correctement, pour un total de 18 points aux temps 1 et 2 et de 24 points au temps 3.

3.2.2.2 Répétition de chiffres

Cette épreuve constitue une mesure typique de la mémoire à court terme (Martinez Perez *et al.*, 2012 ; Swanson et Howell Ashbaker, 2000). D'après les auteurs, elle présente l'avantage d'éliminer toute association sémantique (Chevrie-Muller et Plaza, 2001). Évidemment, cette mesure n'est fiable que si les enfants ont déjà automatisé le

nom des chiffres de 1 à 9. Un point est alloué pour chaque suite de chiffres restituée correctement, et la note de 0 est attribuée dès que la série est incomplète. Si l'enfant échoue dans la répétition d'une série, les séries suivantes se verront toutes attribuer la note de 0. Le résultat total est sur 9 points. Spécifions que le nombre de chiffres dans chaque série va croissant : l'élève est exposé à 3 séquences de 3 chiffres, puis à 3 séquences de 4 chiffres et finalement à 3 séquences de 5 chiffres. Cette épreuve est identique pour les temps 1, 2 et 3.

3.2.2.1 Répétition de rythmes

Cette épreuve permet de mesurer la capacité à décoder, stocker puis rappeler une suite sonore non-verbale (Chevrie-Muller et Plaza, 2001, p. 69). L'élève a pour consigne de reproduire la suite sonore générée par l'expérimentatrice en donnant de légers coups avec son doigt sur une table. Un point est accordé pour chacune des 16 séries de rythmes reproduites de façon adéquate. Cette épreuve a été soumise aux élèves au temps 3 seulement.

3.2.2.3 Répétition de phrases

Cette épreuve est classique dans les batteries servant au dépistage des troubles du langage et de l'articulation (Chevrie-Muller et Plaza, 2001). L'épreuve, telle que proposée dans le NÉEL, consiste à répéter des phrases simples construites de mots très fréquents. Dans leur étude longitudinale étalée sur une période de 7 ans, Butler, Marsh, Sheppard et Sheppard (1985) ont statué que cette épreuve était l'une des plus efficaces dans la prédiction des performances en lecture pendant les premières années de scolarisation. Catts, Fey, Zhang et Tomblin (2001) ont pour leur part observé que l'épreuve de répétition de phrases, évaluée en maternelle, contribuait de façon unique à expliquer la variance dans les résultats en lecture en 2^e année. Étonnamment, les auteurs ont traité l'épreuve comme une mesure du vocabulaire.

Cette ambiguïté dans la classification de l'épreuve de répétition de phrases nous pousse à nous interroger sur les ressources qu'elle mobilise. Les auteurs de la batterie des Nouvelles épreuves pour l'évaluation du langage (Chevrie-Muller et Plaza, 2001) soulignent d'emblée la complexité et la variété des habiletés mises en œuvre lors de la passation de cette épreuve. En effet, en plus des processus de nature cognitive, des processus de natures syntaxique et sémantique sont impliqués. Selon les auteurs (Chevrie-Muller et Plaza, 2001, p. 12), l'enfant mobilise plusieurs habiletés afin de réussir cette épreuve : mémoire à court terme, stockage de segments phonologiques et connaissances lexicales. La tâche est aussi susceptible d'être sensible aux associations sémantiques, à la fréquence, au nombre et à la longueur des mots et à la complexité syntaxique (chunks de mots). La nature de l'épreuve nous oblige à l'assimiler à une épreuve de mémoire, malgré le type d'unités qu'elle met en œuvre.

Cette épreuve a été réalisée aux temps 1 et 2 seulement. L'élève a pour consigne de répéter chacune des 6 phrases, contenant de 3 à 9 mots, tout de suite après son exposition. Les phrases sont de longueurs diverses; la longueur ne va pas croissant. Toutes les phrases de l'épreuve présentent une structure simple de type SV ou SVO. Il est important de le préciser, puisque la clause constitue une unité de stockage importante de la mémoire. Ainsi, une phrase gagne en complexité et induit une charge supplémentaire pour la mémoire, plus le nombre de clauses qu'elle contient est élevé (Roberts et Gibson, 2002). Le NÉEL recommande que le codage de l'épreuve se fasse dans une logique binaire de succès-échec : une erreur dans la répétition de l'énoncé entendu entraîne ainsi la note de zéro pour cet énoncé. Le pointage de l'épreuve s'est fait autrement à Pessamit : chaque mot réussi a été comptabilisé, pour un maximum de 35 points.

3.3 Description des épreuves du temps final

Les épreuves visant à vérifier les performances en lecture des élèves au temps final ont été élaborées dans le cadre de notre recherche, étant donné que les outils disponibles ne convenaient pas parfaitement à nos questions ou à la tranche d'âge étudiée. Les épreuves finales de lecture de mots, de lecture de non-mots et de compréhension en lecture sont donc les seules, parmi celles abordées dans le cadre de ce mémoire, à ne pas être issues de la batterie NÉEL.

3.3.1 Épreuve mesurant la compréhension en lecture

3.3.1.1 Description de l'épreuve de compréhension

Plusieurs études évaluent la compréhension en lecture à l'aide d'une épreuve de lecture silencieuse accompagnée de questions à choix de réponses (Aarnoutse *et al.*, 2001 ; Cain *et al.*, 2004a ; Carretti, Borella, Cornoldi et De Beni, 2009 ; De Beni et Palladino, 2000 ; Foucambert, 2007 ; Swanson et Howell Ashbaker, 2000). Dans l'épreuve standardisée CITO (Staphorsius et Krom, 1998), les questions portent sur plusieurs niveaux du texte – le mot, la phrase et le texte (Netten, Droop et Verhoeven, 2011). Certains auteurs préfèrent que l'épreuve mesure la compréhension de l'implicite du texte. Verhoeven et van Leeuwe (2008), notamment, ont mesuré la compréhension en lecture chez un grand groupe d'enfants, par l'entremise d'une série de textes agrémentés, pour chacun, de 5 questions de compréhension ayant trait à des informations implicites et explicites dans le texte.

La tâche de compréhension en lecture que nous avons conçue consistait à lire 10 textes courts de type documentaire ou narratif et à répondre à la question, en encerclant la meilleure réponse parmi les trois 3 proposées (voir annexe A.3). Les

textes, de tailles dissemblables (de 8 à 25 lignes), étaient présentés en ordre croissant dans un petit livret avec la police Calibri, taille 16. Le texte figure au recto, et la question s'y rapportant, au verso. Les dix textes présentaient des niveaux comparables de difficulté, tant sur le plan du vocabulaire employé (voir le tableau 3.1) que sur le plan de la complexité des phrases. Les participants avaient pour consigne de lire le texte figurant au recto de chacune des pages du cahier, et ce, autant de fois que souhaité, puis de tourner la page et de répondre à la question au verso. Ils devaient y répondre sans retourner au texte, pour éviter toute recherche de mots-clés pouvant avoir une incidence sur les réponses (Simard, Foucambert et Labelle, 2014, p. 52). Chaque question de compréhension comporte une réponse fausse, une question portant sur un élément explicite dans le texte et une question portant sur une information implicite. Chaque choix de réponses, à trois modalités, est précédé de la même formule de présentation : « Choisis la meilleure réponse. Pour toi le texte parle surtout : ».

Afin de nous assurer de la validité de notre jugement quant au pointage de l'épreuve, nous avons demandé à cinq juges d'attribuer une note de 0 à 2 pour chacun des choix de réponses préparés. Les juges que nous avons recrutés sont tous des étudiants de premier cycle de la faculté des sciences humaines de l'Université du Québec à Montréal. Le calcul de fiabilité interjuges pour cette épreuve nous a fourni un alpha de Cronbach de 0,968, ce qui indique qu'elle a un niveau élevé de cohérence.

3.3.1.2 Construction de l'épreuve de compréhension

Les textes de l'épreuve de compréhension en lecture sont tirés des travaux de Foucambert (2003) ainsi que de la version en ligne du magazine *Les Débrouillards*, qui s'adresse à un jeune public. Ils ont été modifiés afin de réduire l'impact de la fréquence et de la complexité syntaxique sur la compréhension (Roberts et Gibson,

2002). Nous avons veillé à ce que les textes soient d'égale complexité sur le plan de la longueur des phrases et de la rareté du vocabulaire utilisé. La longueur des phrases pour chacun des textes a été calculée au moyen du logiciel *Analyse de textes* (Foucambert, 1994). Par la suite, nous avons effectué une ANOVA univariée, ce qui nous a permis de conclure qu'il n'y avait pas de différence significative entre les textes sur le plan de la longueur des phrases ($F(9,85)=0,9$, n.s.). Les textes ont toutefois été classés par ordre de longueur dans le livret, le premier texte faisant 8 lignes et le dernier, 25 lignes. Une analyse de fréquence dans la base *Lextutor* (Cobb, 2013) nous a permis de dresser un portrait général du vocabulaire des 9 textes retenus. La grande majorité des mots utilisés dans les textes proviennent des 1000 mots les plus fréquents en français, comme en témoigne le tableau 3.1. Le tableau montre également l'homogénéité des textes sur le plan du vocabulaire utilisé. Les phrases ont été conçues avec le souci de réduire les ambiguïtés lexicales.

Tableau 3.1 – Portrait de la fréquence du vocabulaire utilisé dans l'épreuve de compréhension en lecture selon la base *Lextutor* (Cobb, 2013)

| Texte | 1-1000 mots plus fréquents | 1001-2000 mots plus fréquents | 2001-3000 mots plus fréquents | Mots hors-liste |
|---------|----------------------------|-------------------------------|-------------------------------|-----------------|
| 1 | 80,3 % | 12,1 % | 3,0 % | 4,6 % |
| 2 | 69,8 % | 4,7 % | 4,7 % | 21,0 % |
| 3 | 81,3 % | 8,3 % | 3,1 % | 7,3 % |
| 4 | 74,1 % | 11,6 % | 0,9 % | 13,4 % |
| 5 | 85,1 % | 5,3 % | 0,9 % | 8,8 % |
| 6 | 76,1 | 9,0 % | 0,8 % | 14,2 % |
| 7 | 71,0 % | 5,8 % | 1,5 % | 21,7 % |
| 8 | 78,6 | 5,6 % | 1,0 % | 14,8 % |
| 9 | 77,3 % | 6,1 % | 1,5 % | 15,2 % |
| Moyenne | 76,4 % | 6,9 % | 1,7 % | 15,0 % |

Dans la mesure où nous cherchions à travailler sur les relations que l'élève peut établir afin d'établir une compréhension cohérente du texte lu, nous avons construit les questions de façon à ce que, parmi les trois choix soumis pour chaque texte, on

retrouve toujours une réponse fausse, une réponse adéquate et une réponse dénotant une compréhension accrue du texte. Une grande place a été accordée à la notion d'inférence dans la conception de l'épreuve. La notion d'inférence renvoie à la production d'informations nécessaires à la compréhension qui n'ont pas été formulées de façon explicite (Rossi, 2008). Ainsi, deux points étaient attribués si l'élève sélectionnait la réponse dénotant une compréhension plus sophistiquée du texte, soit celle qui était constituée à partir des inférences qu'on pouvait dégager de celui-ci. Un point était accordé pour une réponse adéquate, reprenant un élément explicite du texte, le participant démontré une compréhension superficielle du texte. Aucun point n'était attribué pour avoir choisi la réponse fausse.

3.3.1.3 Codage de l'épreuve de compréhension

Les données de l'épreuve de compréhension ont mis en évidence une disparité dans la réussite des textes. Au regard de cette observation, nous avons donc décidé d'affiner le codage de l'épreuve a posteriori, afin de mieux représenter l'effort qui devait être déployé par les élèves pour obtenir un bon score à l'épreuve.

Notre premier raisonnement porte sur les textes. Il est raisonnable de conclure qu'un texte n'ayant recueilli que peu de bonnes réponses, soit la note de 2, est probablement plus difficile qu'un texte ayant recueilli un grand nombre de bonnes réponses. Ainsi, nous avons attribué plus de points à une bonne réponse fournie à une question portant sur un texte difficile. La difficulté des textes se traduisait par la grande variation entre les taux de réussite des élèves – soit les réponses notées 2 – à travers les textes. Le tableau 3.2 illustre d'ailleurs cette variation, où on retrouve une somme des bonnes réponses pour chaque texte.

Tableau 3.2 – Pointage accumulé pour les bonnes réponses (notées sur 2) pour chaque texte de l'épreuve de compréhension en lecture (n=71)

| Numéro du texte | Bonnes réponses |
|-----------------|-----------------|
| Texte 1 | 122 points |
| Texte 2 | 110 points |
| Textes 10 et 7 | 98 points |
| Textes 4 et 6 | 80 points |
| Texte 9 | 72 points |
| Texte 8 | 60 points |
| Texte 5 | 48 points |
| Texte 3 | 44 points |

Nous avons décidé de considérer comme difficile un texte étant de 1,96 fois l'écart-type (30,25 points) supérieur à la moyenne (81,2 points). Ce calcul nous a permis de conclure que les textes 3, 5 et 8 devaient être considérés comme difficiles. Les bonnes réponses fournies pour ces deux textes se sont vues accorder une note de 4 points, au lieu de 2. Cette modification de codage est légitimée par la difficulté inhérente aux deux textes : les élèves ayant bien répondu à ces textes témoignent d'une acuité supplémentaire par rapport à ce qui est attendu dans les autres textes, et ceci doit se traduire par une note plus élevée. Il est à noter que les autres textes n'ont pas bénéficié d'un traitement particulier. Cette modification vise donc ultimement à respecter le caractère complexe de la compréhension en lecture et l'hétérogénéité des textes sur le plan de leur difficulté.

Toutefois, nous voulions nous assurer de bien identifier les textes représentant une réelle difficulté, et non de donner plus de poids à un item dysfonctionnel. Ainsi, notre deuxième raisonnement, qui vise à sélectionner les textes ayant le mieux discriminé les bons compreneurs, porte sur les individus. Il est attendu qu'une épreuve puisse discriminer entre les forts et les faibles. Dans le but de repérer les textes les plus discriminants, nous avons fait une analyse visant à comparer la performance des élèves par rapport aux résultats globaux, pour tous les textes. Les calculs qui nous ont

permis, d'une part, de modifier notre le codage des textes difficiles, puis, d'autre part, de vérifier la pertinence de ce codage, se retrouvent dans les paragraphes qui suivent.

Nous avons divisé les élèves en trois groupes : les forts, les faibles et les moyens. Les élèves ayant de 17 à 20 points constituent le groupe fort ($n=17$), ceux ayant cumulé de 14 à 16 points faisaient partie du groupe moyen ($n=23$), et ceux ayant eu un résultat de moins de 14 points sur 20 ont été classés dans le groupe faible ($n=25$). En calculant le khi-deux entre les résultats globaux des groupes et les résultats pour chaque texte, il s'est avéré que les textes 5 et 3 discriminaient significativement les groupes. Le texte 8, quant à lui, ne permettait pas de bien discriminer entre les groupes. Les résultats incohérents des participants, mis en exergue par les résultats des analyses du khi-deux, nous ont incitée à considérer le texte 8 comme un item dysfonctionnel. Nous avons donc exclu ce texte des analyses ultérieures, dans le but d'éviter de donner un plus grand poids à un item problématique.

Les bonnes réponses des textes 3 et 5 ont ainsi bénéficié du double de points (7 textes X 2 points ; 2 textes X 4 points). Les modifications apportées au codage portent donc à 22 points le total pour l'épreuve de compréhension de textes et à 9 le nombre de textes par participant. Par rapport aux données initiales, le nouveau codage a recentré les données. En effet, alors que la moyenne était de 74% initialement (avec un écart-type de 13%), elle se situe plutôt à 67,91% (avec un écart-type de 13,1%).

3.3.2 Épreuves mesurant le décodage

Deux mesures ont été utilisées pour évaluer les performances des élèves en décodage. Il s'agit de mesures classiques de lecture de non-mots et de lecture de mots. Selon Tunmer et Chapman (2012), l'utilisation d'un score composite de lecture de mots et de non-mots constitue la stratégie idéale pour évaluer le décodage.

3.3.2.1 Lecture de non-mots en une minute

L'épreuve de lecture de non-mots en une minute est une méthode efficace et éprouvée pour mesurer les performances en décodage, sans que l'élève puisse se référer à un support sémantique (Martinez Perez *et al.*, 2012). Tunmer et Chapman (2012) ont d'ailleurs utilisé une tâche équivalente de lecture de non-mots en contrôlant autrement la variable temps.

Dans notre épreuve, l'élève doit lire à voix haute une série de 77 non-mots dont la structure syllabique et le nombre de syllabes varient. Les non-mots proviennent de différentes tâches de la batterie BALE (2010) et respectent les contraintes phonotactiques du français. On demande à l'élève de lire les mots qui sont sur la feuille le mieux et le plus vite possible, en l'avertissant que ce ne sont pas des mots français. La minute allouée à la passation de l'épreuve est calculée à l'aide d'un chronomètre. Le taux de non-mots lus correctement a été comptabilisé pour constituer le résultat final de chaque élève. Les consignes et la liste de mots utilisés figurent à l'annexe A.3.

3.3.2.2 Lecture de mots en une minute

L'épreuve de lecture de mots en une minute est également utilisée dans plusieurs études qui traitent du décodage (Aarnoutse *et al.*, 2001 ; Carretti *et al.*, 2009 ; Verhoeven et van Leeuwe, 2008). Dans cette épreuve, l'enfant doit lire, à voix haute, une liste de 104 mots dont le nombre de syllabes va croissant. Le taux de mots correctement lus par minute constitue le résultat pour chaque élève. La minute allouée à la passation de l'épreuve est calculée à l'aide d'un chronomètre. Notre épreuve est tirée de l'épreuve d'évaluation de la compétence en lecture – Lecture de Mots et

Compréhension – Révisée conçue par Khomsi (1999). Les consignes et la liste des mots se trouvent à l'annexe A.2.

3.3.3 Mise à l'essai

Nous avons effectué des tests préliminaires dans une classe de 3^e année d'une école primaire de la Commission scolaire de Saint-Hyacinthe pour nous assurer du bon fonctionnement des tâches de décodage et de compréhension. Cette mise à l'essai avait d'abord pour but de valider les procédures de passation des épreuves (temps alloué pour chaque tâche, difficulté des épreuves, etc.) De plus, nous voulions vérifier que le vocabulaire des épreuves de lecture de mots et de non-mots ne présentait pas de problème particulier. Par ailleurs, nous voulions confirmer la validité des items qui seraient présentés lors de l'étude principale. Pour finir, nous voulions vérifier que l'épreuve de compréhension de textes ne présentait pas d'effet plafond ou plancher.

Suite à ces essais, certains items de l'épreuve de lecture de mots ont été éliminés, étant donné qu'ils semblaient plus difficiles que le reste des items, probablement à cause de leur rareté. Le terme «fenouil» faisait partie de cette catégorie. Quelques items de l'épreuve de lecture de non-mots se sont vus retirés parce qu'ils ressemblaient de façon trop frappante à un mot véritable et menaçaient de fausser les résultats (notamment « aivron », fort similaire à *aviron*).

L'épreuve de compréhension a généré une grande variabilité dans les résultats, ce que nous espérions, et nous avons décidé de la présenter telle quelle aux participants de l'étude principale.

3.4 Synthèse des variables et des épreuves

Au total, 18 épreuves aux temps 1, 2 et 3 (6 épreuves par temps) et 3 épreuves en temps final ont été utilisées pour répondre à nos questions de recherche. Le tableau 3.3 présente une synthèse de chacun des instruments de mesures et de l'habileté que nous souhaitons évaluer à l'aide de cet instrument. Nous y indiquons aussi la modalité de la réponse et le nombre d'items associés à chacune des épreuves. Ce type d'informations nous permet de légitimer les regroupements de sous-tests auxquels nous avons procédé afin de simplifier nos analyses statistiques, qui feront l'objet du chapitre IV. Par exemple, les sous-tests Vocabulaire 1 et Vocabulaire 2 ont été rassemblés sous le libellé « Vocabulaire (réception) » dans la présente étude, étant donné que les modalités de réponses et la nature des processus mobilisés étaient semblables. Le tableau synthétique 4.3.3 présente aussi l'avantage d'aider à expliquer ce qui ressortira de l'analyse des données. En effet, les tâches ne sont pas « pures », et différentes habiletés peuvent être mises à contribution selon la nature de la tâche.

Tableau 3.3 – Synthèse des habiletés latentes anticipées et des épreuves

| Habilitété | Épreuves | Items | Nature de la réponse | Point alloués |
|----------------------------------|---------------------------------|-------------|---|------------------|
| Temps 1 et Temps 2 | | | | |
| Mémoire à court terme | Répétition de mots rares | 2 x 9 mots | Orale | 18 points |
| | Répétition de chiffres | 3 suites | Orale | 9 points |
| | Répétition de phrases | 6 phrases | Orale | 35 points |
| Topologie et arithmétique | | | | |
| | Topologie et arithmétique | 7 actions | Manipulation d'une figurine de chat | 12 points |
| | Arithmétique 1 | 5 actions | Manipulation de figurines d'animaux, désignation et réponse orale | 10 points |
| Vocabulaire | Arithmétique 2 | 4 actions | Désignation | 4 points |
| | Vocabulaire (expression) | | | 68 points |
| | Vocabulaire 1 | 14 images | Orale | 28 points |
| | Vocabulaire 2 | 20 images | Orale | 40 points |
| Vocabulaire (réception) | | | | |
| | Vocabulaire 1 | 36 images | Désignation de l'image | 36 points |
| | Vocabulaire 2 | 20 images | Désignation de l'image | 20 points |
| Temps 3 | | | | |
| Mémoire à court terme | Répétition de mots rares | 2 x 12 mots | Orale | 24 points |
| | Répétition de chiffres | 4 suites | Orale | 9 points |
| | Répétition de rythmes | 16 suites | Suite sonore non-verbale | 16 points |

| Habilité | Épreuves | Items | Nature de la réponse | Point alloués |
|--------------------|--|-----------------------|--|-------------------|
| | Topologie et arithmétique | | | 30 points |
| | Topologie 2 | 7 actions | Manipulation de figurines de canards | 14 points |
| | Topologie 3 | 8 actions | Manipulation de figurines d'animaux et désignation | 16 points |
| Vocabulaire | Vocabulaire (expression) | | | 112 points |
| | Vocabulaire 1 | 36 images | Orale | 72 points |
| | Vocabulaire 2 | 20 images | Orale | 40 points |
| | Vocabulaire (réception) | | | 56 points |
| | Vocabulaire 1 | 36 images | Désignation de l'image | 36 points |
| | Vocabulaire 2 | 20/21 images | Désignation de l'image | 20 points |
| Temps final | | | | |
| Compréhension | Compréhension de textes courts | 9 textes, 9 questions | Sélection d'une réponse parmi 3 choix proposés | 22 points |
| Décodage | Lecture de non-mots en une minute | 77 non-mots | Orale | 77 points |
| | Lecture de mots en une minute | 104 mots | Orale | 104 points |

3.5 Traitement des données

Dès le début du codage, chaque participant s'est vu accorder un code allant de 1 à 76 afin d'assurer l'anonymat des données. Les résultats de chaque épreuve ont tous été ramenés sur 1, pour faciliter le traitement des données. Afin de répondre à nos deux grandes questions de recherche, nous avons utilisé différentes analyses statistiques. Nous avons tout d'abord vérifié la normalité des distributions de nos résultats pour chaque épreuve et pour tous les temps (temps 1, temps 2, temps 3 et temps final). Le procédé de normalisation des données constitue un préalable à l'utilisation des statistiques paramétriques auxquelles nous souhaitons soumettre nos données.

Dans un premier temps, nous nous sommes penchée sur notre objectif de nature diagnostique. Nous avons d'abord procédé aux analyses des corrélations entre nos variables mesurées. Ce procédé permet d'approximer la force du lien unissant certaines variables et de dépister les corrélations surprenantes. De plus, l'analyse des corrélations constitue la première étape vers les analyses de régression que nous avons effectuées par la suite, avec nos variables mesurées. L'analyse de régression permet de construire un modèle pertinent à l'aide de variables prédictives non redondantes. En effet, Larson-Hall (2012, p. 184)(cité dans Fortier, 2013, p. 99) indique que la multicolinéarité est un phénomène à identifier avant de procéder à des analyses de régression, puisque la présence d'intercorrélations trop fortes ($> 0,8$) entre variables prédictives peut mener à des résultats erronés. Nous avons donc effectué des analyses de régression linéaire multiple afin de déterminer lesquelles, parmi les épreuves aux temps 1, 2 et 3, constituaient les meilleurs prédicteurs des performances en lecture, en compréhension et en décodage au temps final. Ces analyses nous ont permis de répondre à notre première grande question de recherche, dont le mandat était de déterminer quelles épreuves étaient les plus efficaces pour diagnostiquer des problèmes ultérieurs dans les performances en lecture.

Dans un deuxième temps, nous nous sommes intéressée à notre objectif de nature explicative. Ainsi, nous avons soumis nos données à des analyses en composantes principales (ACP). Ce type d'analyses crée des composantes, sous forme d'axes, par un travail de synthétisation de l'information, sans que le chercheur ne fasse de regroupements a priori. L'analyse a mis en exergue deux axes ou facteurs récurrents à travers les temps 1, 2 et 3. Les axes construits de nos ACP correspondent à nos variables latentes. Ce procédé nous a permis de légitimer les regroupements que nous avons établis afin de construire notre modèle de régression final, impliquant nos variables latentes MCT et Vocabulaire. Nous avons, préalablement à cette manipulation statistique, effectué les analyses des corrélations d'usage entre nos variables latentes.

Nous avons soumis nos variables construites par ACP à des analyses de régression. Celles-ci nous ont permis de mesurer la contribution relative de chacune des variables aux modèles des temps 1, 2 et 3. Nous avons ainsi pu évaluer l'influence de notre variable langagière, le vocabulaire, et de notre variable cognitive, la mémoire, sur les performances ultérieures en décodage et en compréhension, afin de répondre à notre deuxième grande question de recherche.

CHAPITRE IV

RÉSULTATS

Ce chapitre est divisé en deux parties, chacune représentant l'un de nos deux grands objectifs de recherche. La section 4.1 traite des analyses appliquées aux variables mesurées ou brutes. Cette section fournit les résultats nécessaires à l'analyse de notre objectif de nature diagnostique. La section 4.2 présente les résultats associés aux analyses des variables latentes, construites par analyses factorielles en composantes principales, et permet de répondre à notre objectif de nature explicative. Dans chacune de ces parties, nous présenterons les données sous un même gabarit. Nous commençons par la présentation des statistiques descriptives que nous avons utilisées afin de faire un portrait des données à analyser. Par la suite, nous exposons l'ensemble des résultats pour chacune des analyses effectuées, soit les analyses de corrélation, les analyses de régression et les analyses en composantes principales.

4.1 Analyses appliquées aux variables mesurées

4.1.1 Statistiques descriptives pour les variables mesurées

Nous avons procédé à des analyses descriptives de nos données afin de vérifier la normalité des distributions pour chaque variable, pour tous les temps (temps 1, 2, 3, et final). Rappelons que les variables dites « mesurées » correspondent aux résultats des épreuves pour les temps 1, 2 et 3. Ce sont ces mesures qui nous permettront de

répondre adéquatement à notre première question de recherche, de nature diagnostique. Les coefficients d'asymétrie et d'aplatissement ont d'abord été calculés. La pratique suggère que des valeurs de symétrie et des valeurs d'aplatissement allant de -1,5 à 1,5 nous permettent de considérer que la variable présente une distribution normale, suivant ce que recommandent Carricano, Poujol et Bertrandias (2010, p. 35). Nous avons par la suite examiné les représentations graphiques de chacune des distributions, ce qui nous a permis de confirmer nos précédents jugements de normalité. À la suite de cette série d'observations, nous avons constaté que 8 variables présentaient une distribution anormale. Nous avons donc procédé à une transformation de ces variables en les élevant au carré. Ces manipulations ont permis la normalisation de l'ensemble des distributions, une condition nécessaire à l'utilisation des statistiques paramétriques qui s'ensuivent (Carricano *et al.*, 2010, p. 35). Le tableau suivant indique les données qui ont été utilisées pour faire les analyses, ainsi que les transformations opérées, s'il y a lieu. Les coefficients d'asymétrie et d'aplatissement, utilisés pour vérifier la normalité des distributions, sont aussi indiqués.

Tableau 4.1 – Statistiques descriptives des variables à l'étude retenues pour l'analyse (après transformations)

| Épreuves | Min | Max | N | Moy | Écart type | Asymétrie | Aplatissement | Transformation |
|---------------------------------------|-------|-------|----|------|------------|-----------|---------------|----------------|
| Répétition de mots rares | 0,11 | 1,00 | 50 | 0,61 | 0,22 | -0,13 | -0,83 | aucune |
| Répétition de chiffres | 0,00 | 0,89 | 50 | 0,45 | 0,22 | 0,19 | -0,82 | aucune |
| Topologie et arithmétique | 0,12 | 1,00 | 50 | 0,59 | 0,20 | -0,22 | 0,17 | aucune |
| ⌊ Vocabulaire (expression) | 0,35 | 0,90 | 50 | 0,64 | 0,13 | 0,00 | -0,24 | aucune |
| Répétition de phrases | 0,10 | 1,00 | 50 | 0,65 | 0,21 | -0,30 | 0,02 | au carré |
| Vocabulaire (réception) | 0,46 | 0,91 | 50 | 0,72 | 0,11 | -0,42 | -0,60 | aucune |
| Vocabulaire (expression et réception) | 0,46 | 0,90 | 50 | 0,68 | 0,11 | 0,12 | -0,68 | aucune |
| Répétition de mots rares | 0,17 | 0,75 | 48 | 0,47 | 0,13 | 0,17 | 0,02 | aucune |
| Répétition de chiffres | 0,00 | 1,00 | 48 | 0,57 | 0,23 | -0,35 | -0,12 | aucune |
| Topologie et arithmétique | 0,04 | 1,00 | 48 | 0,71 | 0,26 | -0,84 | -0,07 | au carré |
| ⌊ Vocabulaire (expression) | 0,28 | 0,89 | 48 | 0,61 | 0,15 | -0,05 | -0,89 | au carré |
| Répétition de phrases | 0,43 | 1,00 | 48 | 0,76 | 0,15 | -0,05 | -0,84 | au carré |
| Vocabulaire (réception) | 0,20 | 0,93 | 48 | 0,65 | 0,16 | -0,71 | 0,55 | au carré |
| Vocabulaire (expression et réception) | 0,24 | 0,89 | 48 | 0,63 | 0,14 | -0,29 | 0,05 | aucune |
| Répétition de mots rares | 0,21 | 0,96 | 53 | 0,61 | 0,20 | -0,05 | -0,91 | aucune |
| Répétition de chiffres | 0,05 | 1,00 | 53 | 0,64 | 0,24 | -0,08 | -0,51 | au carré |
| Répétition de rythmes | 0,31 | 1,00 | 53 | 0,73 | 0,17 | -0,43 | -0,33 | aucune |
| ⌊ Topologie et arithmétique | 0,07 | 1,00 | 53 | 0,66 | 0,25 | -0,27 | -0,60 | au carré |
| Vocabulaire (expression) | 0,45 | 0,90 | 53 | 0,70 | 0,10 | -0,27 | -0,47 | aucune |
| Vocabulaire (réception) | 0,24 | 1,00 | 53 | 0,73 | 0,17 | -0,72 | 0,12 | au carré |
| Vocabulaire (expression et réception) | 0,35 | 0,95 | 53 | 0,71 | 0,13 | -0,61 | 0,20 | aucune |
| Compréhension de textes | -0,10 | -0,15 | 70 | 0,69 | 0,16 | 0,32 | 1,00 | aucune |
| ⌊ Lecture de mots en une minute | 0,24 | 0,96 | 70 | 0,61 | 0,18 | -0,03 | -0,93 | aucune |
| Lecture de non-mots en une minute | 0,59 | -0,55 | 70 | 0,45 | 0,20 | 0,13 | 0,89 | aucune |

4.1.2 Analyses de corrélation

Avant de procéder à l'évaluation des liens unissant les épreuves des temps 1, 2 et 3 avec les épreuves de lecture en temps final au moyen des analyses de régression, nous avons réalisé des analyses de corrélation, qui feront l'objet de la présente section. Les tableaux 4.2, 4.3 et 4.4 présentent les valeurs des coefficients de corrélation (r) de Pearson découlant des analyses de corrélation pour chacun des temps. Les épreuves des temps 1, 2 et 3 sont mises en relation avec celles du temps final.

Tableau 4.2 – Corrélations entre les variables du T1 et du TF

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------------------------------|---------------|---------------|---------------|---------------|--------------|-------------|---------------|---------------|---------|
| 1. Répétition de mots rares | r n | | | | | | | | |
| | 0,466** 50 | 1 50 | | | | | | | |
| 2. Répétition de chiffre | r n | | | | | | | | |
| | 0,181 50 | 0,261 50 | 1 50 | | | | | | |
| 3. Topologie et arithmétique | r n | | | | | | | | |
| | 0,213 50 | 0,275 50 | 0,507** 50 | 1 50 | | | | | |
| 4. Vocabulaire (expression) | r n | | | | | | | | |
| | 0,595** 50 | 0,584** 50 | 0,376** 50 | 0,450** 50 | 1 50 | | | | |
| 5. Répétition de phrases | r n | | | | | | | | |
| | 0,142 50 | 0,307* 50 | 0,399** 50 | 0,620** 50 | 0,322* 50 | 1 50 | | | |
| 6. Vocabulaire (réception) | r n | | | | | | | | |
| | -0,145 48 | 0,146 48 | -0,013 48 | 0,089 48 | 0,133 48 | 0,125 48 | 1 70 | | |
| 7. Lecture de mots | r n | | | | | | | | |
| | -0,056 48 | 0,126 48 | 0,048 48 | 0,068 48 | 0,106 48 | 0,057 48 | 0,820** 70 | 1 70 | |
| 8. Lecture de non-mots | r n | | | | | | | | |
| | -0,021 48 | 0,227 48 | 0,095 48 | 0,311* 48 | 0,310* 48 | 0,19 48 | 0,359** 70 | 0,327** 70 | 1 70 |
| 9. Compréhension | r n | | | | | | | | |

Note 1 : n= 50; * p<0,05; ** p<0,01

Note 2 : Les traits visent à faciliter la lecture du tableau en contrastant les épreuves en temps 1, 2 et 3 des épreuves finales.

4.1.2.1 Corrélations au temps 1

Il est intéressant d'observer que les trois épreuves de répétition sont significativement corrélées. En effet, la répétition de chiffres, la répétition de mots rares et la répétition de phrases présentent des corrélations positives plutôt fortes (r de 0,47 à 0,60; $p < 0,01$). Cette observation tend à appuyer l'une de nos hypothèses de recherche, qui stipule que les épreuves de répétition circonscrivent une habileté plus large, que la littérature désigne comme la mémoire à court terme. Les analyses factorielles (4.2.1) nous permettront de confirmer ces analyses préliminaires. Un même type de corrélation se présente chez les épreuves représentant le vocabulaire. En effet, il appert que les épreuves d'expression, de réception et de topologie et arithmétique présentent des corrélations plutôt fortes (r de 0,39 à 0,62; $p < 0,01$). Par ailleurs, l'épreuve de répétition de phrases est la seule qui soit corrélée significativement ($p < 0,01$) avec toutes les autres épreuves du temps 1. Cette observation témoigne du caractère composite de cette épreuve, qui semble mobiliser plusieurs habiletés plutôt que d'en représenter une indépendamment. Il est étonnant de constater la faiblesse des corrélations entre les épreuves du temps 1 et les épreuves de lecture en temps final : seules les épreuves de vocabulaire (expression) ($r=0,31$; $p < 0,05$) et de répétition de phrases ($r=0,31$; $p < 0,05$) se sont avérées corrélées avec l'épreuve de compréhension. Soulignons que nos variables calculées finales qui représentent les performances en décodage ne sont corrélées avec aucune variable du temps 1.

4.1.2.2 Corrélations au temps 2

D'entrée de jeu, mentionnons qu'une fois de plus, l'épreuve de répétition de phrases et les autres épreuves du temps 2 (r de 0,42 à 0,62; $p < 0,01$) entretiennent des corrélations significatives plutôt fortes. Parmi les épreuves initiales, seule celle de vocabulaire (réception) est faiblement corrélée avec une variable finale, soit l'épreuve

de compréhension de textes (r de 0,29; $p < 0,05$). Globalement, les résultats du temps 2 exhibent des tendances similaires à celles observées au temps 1 : les épreuves de répétition sont corrélées entre elles (r de 0,51 à 0,57; $p < 0,01$) et les épreuves de vocabulaire en font de même (r de 0,50 à 0,67; $p < 0,01$).

4.1.2.3 Corrélations au temps 3

Les épreuves du temps 3 voient les patrons des temps 1 et 2 se confirmer. En effet, les épreuves de répétition sont corrélées de modérément à fortement (r de 0,40 à 0,79; $p < 0,01$). Rappelons que l'épreuve de répétition de phrases n'est pas disponible au temps 3. Toutefois, une épreuve de répétition de rythmes est introduite. De plus, les épreuves de vocabulaire sont plus fortement corrélées qu'aux deux temps précédents (r de 0,50 à 0,80; $p < 0,01$). Par ailleurs, de nouvelles corrélations significatives émergent entre les épreuves : les grandes habiletés qui étaient révélées par les analyses de corrélation antérieures semblent se confirmer au temps 3. Pour finir, il est intéressant d'observer que les épreuves vocabulaire (expression), vocabulaire (réception) et topologie et arithmétique sont corrélées avec l'épreuve finale de compréhension.

Tableau 4.3 – Corrélations entre toutes les variables (Temps 2)

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------------------------------|-------------------|---------------|---------------|---------------|---------------|--------------|---------------|---------------|---------|
| 1. Répétition de mots rares | r 1 n 49 | | | | | | | | |
| 2. Répétition de chiffre | r 0,533** n 48 | 1 48 | | | | | | | |
| 3. Topologie et arithmétique | r 0,296* n 49 | 0,193 48 | 1 49 | | | | | | |
| 4. Vocabulaire (expression). | r 0,267 n 49 | 0,216 48 | 0,521** 49 | 1 49 | | | | | |
| 5. Répétition de phrases | r 0,572** n 49 | 0,509** 48 | 0,422** 49 | 0,463** 49 | 1 49 | | | | |
| 6. Vocabulaire (réception) | r 0,375** n 49 | 0,11 48 | 0,505** 49 | 0,673** 49 | 0,617** 49 | 1 49 | | | |
| 7. Lecture de mots | r -0,006 n 47 | -0,011 46 | -0,154 47 | -0,064 47 | 0,107 47 | 0,073 47 | 1 70 | | |
| 8. Lecture de non-mots | r -0,18 n 47 | -0,107 46 | -0,158 47 | -0,126 47 | -0,055 47 | 0,001 47 | 0,820** 70 | 1 70 | |
| 9. Compréhension | r -0,018 n 47 | 0,033 46 | -0,059 47 | 0,164 47 | 0,19 47 | 0,288* 47 | 0,359** 70 | 0,327** 70 | 1 70 |

Note 1 : n= 49; * p<0,05; ** p<0,01

Note 2 : Les traits visent à faciliter la lecture du tableau en contrastant les épreuves en temps 1, 2 et 3 des épreuves finales.

Tableau 4.4 – Corrélations entre toutes les variables (Temps 3)

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------------------------------|-------------------|-----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|---------|
| 1. Répétition de mots rares | r 1 n 54 | | | | | | | | |
| 2. Répétition de chiffres | r 0,793** n 54 | 1 54 | | | | | | | |
| 3. Répétition de rythmes | r 0,396** n 54 | r 0,479** 54 | 1 54 | | | | | | |
| 4. Topologie et arithmétique | r 0,471** n 54 | r 0,397** 54 | r 0,318* 54 | 1 55 | | | | | |
| 5. Vocabulaire (expression) | r 0,363** n 54 | r 0,314* 54 | r 0,079 54 | r 0,503** 54 | 1 54 | | | | |
| 6. Vocabulaire (réception) | r 0,446** n 53 | r 0,392** 53 | r 0,206 53 | r 0,604** 53 | r 0,800** 53 | 1 53 | | | |
| 7. Lecture de mots | r -0,021 n 49 | r 0,101 49 | r 0,09 49 | r -0,023 50 | r 0,184 49 | r 0,202 48 | 1 70 | | |
| 8. Lecture de non-mots | r -0,047 n 49 | r 0,074 49 | r 0,045 49 | r -0,127 50 | r 0,144 49 | r 0,125 48 | r 0,820** 70 | 1 70 | |
| 9. Compréhension | r 0,258 n 49 | r 0,21 49 | r 0,017 49 | r 0,286* 50 | r 0,526** 49 | r 0,521** 48 | r 0,359** 70 | r 0,327** 70 | 1 70 |

Note 1 : n= 54; * p<0,05; ** p<0,01

Note 2 : Les traits visent à faciliter la lecture du tableau en contrastant les épreuves en temps 1, 2 et 3 des épreuves finales.

4.1.2.4 Multicolinéarité et modifications

Nous avons analysé les corrélations présentes entre les variables prédictives, soit celles qui sont présentes entre les épreuves des temps 1, 2 et 3, afin d'identifier celles pouvant présenter un lien de multicolinéarité. Ce terme renvoie à des intercorrélations des variables indépendantes contenant des informations similaires ou qui se chevauchent. L'inclusion de ces variables dans un modèle de régression peut affaiblir la significativité des coefficients de régression et renforcer les erreurs de type 1, qui correspondent à un rejet d'une hypothèse nulle vraie (Howell, 2008, p. 98). Comme il est possible de le constater en observant les tableaux 4.2., 4.3 et 4.4, les seules variables s'étant avérées trop fortement intercorrélées sont les variables vocabulaire (expression) et vocabulaire (réception) au temps 3 ($r= 0,80$, $p < 0,01$). Dans le but de les utiliser dans l'analyse de régression, qui fait l'objet de la section 4.1.3, nous avons jumelé ces deux variables dans les trois temps. Ces variables présentaient des corrélations plutôt fortes aux temps 1, 2 et 3 et partageaient aussi un certain nombre d'items-cibles (voir le chapitre III sur la méthode). Par ailleurs, en plus de bénéficier d'*a priori* théoriques, le couplage de ces deux épreuves se justifie par les limites que nous imposent les analyses de régression. En effet, un modèle de régression efficace se doit de comporter globalement une variable par 10 participants (Wilson VanVoorhis et Morgan, 2007, p. 48). Notre échantillon étant réduit, nous nous sommes fixé une limite de 5 variables prédictives par modèle.

4.1.3 Analyses de régression

La présente section nous permettra de répondre à notre première question de recherche, qui vise à déterminer les mesures qui prédisent le mieux les performances ultérieures en lecture. Pour arriver à nos fins, nous avons effectué des analyses de régression. Ces analyses permettent de vérifier quels sont les meilleurs prédicteurs

d'une variable expliquée. En effet, elles mesurent la portée explicative de l'ensemble des variables sur la variable finale. Les valeurs bêta nous permettent de comparer le poids respectif de chaque variable au sein de chaque modèle. Dans le présent travail de recherche, nous avons utilisé deux variables finales : les performances en décodage, qui est une mesure composite, et les performances en compréhension, qui est une variable mesurée. Les différentes analyses effectuées portent donc sur deux sous-ensembles d'analyses : les variables prédictives T1, T2 et T3 avec le décodage en temps final, puis les variables prédictives T1, T2 et T3 avec la compréhension en temps final.

Les variables prédictives qui ont été utilisées dans les analyses de régression pour parvenir aux analyses de régression aux temps 1 et 2 sont les suivantes : Répétition de mots rares, Répétition de chiffres, Topologie et arithmétique 1 et 2, Répétition de phrases, Vocabulaire (réception et expression). Les variables prédictives au temps 3 diffèrent quelque peu de celles des temps 1 et 2 (se référer au chapitre 3 pour les détails). En voici la liste : Répétition de mots rares, Répétition de chiffres, Répétition de rythmes, Topologie 2 et 3, Répétition de phrases, Vocabulaire (réception et expression). Le tableau 4.5 présente les modèles de régression qui ont été testés à l'aide des variables prédictives citées ci-haut et des deux variables prédites, soit le décodage et la compréhension. Il présente le premier modèle qui a été généré par l'analyse de régression, suivi du deuxième, avec les valeurs se rapportant à chacun. En effet, comme il est d'usage quand on cherche à prédire, nous avons systématiquement procédé à l'analyse des résidus à la suite de chaque premier modèle (Cohen, Cohen, West et Aiken, 2003). Les résidus correspondent à l'erreur entre la prédiction faite par le modèle et la mesure réelle pour chaque participant. Les individus qui étaient mal classés par les analyses – c'est-à-dire les individus présentant des résultats supérieurs à $\pm 1,96$ écart-type de la moyenne des résidus – ont été exclus, afin de générer un deuxième modèle s'ajustant mieux aux données. Un maximum de 6,5% des individus ont ainsi été écartés à la suite de l'analyse des

résidus, soit un maximum de trois participants par modèle. Il est à noter les données indiquées à droite dans le tableau 4.5 représentent ceux des modèles finaux, après élimination des résidus.

Tableau 4.5 – Six modèles de régression testés (après ajustements)

| Variable prédite | Modèle 1 | | | Modèle 2 (après ajustements) | | | | | | |
|------------------|----------|----------------|-------|------------------------------|----------------|-------|----|--------------|---------|-----|
| | R | R ² | n | R | R ² | n | F | P | Gain | |
| Décodage | T1 | 0,300 | 0,09 | 48 | 0,433 | 0,188 | 46 | (5,45)=1,849 | 0,125 | 10% |
| | T2 | 0,204 | 0,042 | 46 | 0,250 | 0,063 | 45 | (5,44)=0,522 | 0,758 | 2% |
| | T3 | 0,365 | 0,133 | 48 | 0,520 | 0,271 | 46 | (5,45)=2,972 | 0,023* | 14% |
| Compréhension | T1 | 0,453 | 0,205 | 48 | 0,615 | 0,378 | 45 | (5,44)=4,737 | 0,002** | 17% |
| | T2 | 0,377 | 0,142 | 46 | 0,538 | 0,29 | 43 | (5,42)=3,02 | 0,022* | 15% |
| | T3 | 0,553 | 0,306 | 48 | 0,703 | 0,495 | 45 | (5,45)=7,838 | 0** | 19% |

Rappelons que les résultats exprimés dans le tableau ne proviennent pas d'exactly les mêmes enfants aux trois temps. Les analyses de régression tiennent compte de cette réalité. En guise d'éclaircissement, dans le modèle mettant en relation la compréhension au temps final et les variables du temps 1, qui est en gris dans le tableau 4.5, les résultats sont générés à partir de résultats de 48 élèves dont nous aurons mesuré la compréhension au temps final et qui auront passé l'ensemble des épreuves du temps 1.

Malgré les modifications effectuées grâce à l'analyse des résidus, deux modèles se sont avérés non-significatifs. Nous considérerons seulement les coefficients des 4 modèles qui se sont avérés significatifs, sur les six modèles testés. Les données du tableau 4.5 permettent néanmoins de constater que la création de modèles après transformations a permis des gains en explication de la variance. La dernière colonne indique le gain en explication de la variance que l'élimination de résidus nous a permis d'obtenir entre le premier et le deuxième modèle.

Nous souhaitons ajouter, à titre de justification préventive, que le R^2 explique un pourcentage de la variance dans les résultats de la variable finale, et que le fait qu'il soit petit ne le rend pas inintéressant. Dans la mesure où le phénomène étudié est complexe et influencé par une panoplie de facteurs (Fayol, 1992), il est normal qu'il soit difficile de l'expliquer. La portion de la variance qui est expliquée – c'est-à-dire le R^2 – a tendance à être plus grande lorsque des variables individuelles, comme la scolarité des parents ou la motivation générale envers l'école, sont introduites dans le modèle. Par ailleurs, les R^2 que nous avons obtenus sont comparables à ceux observés dans d'autres études. Par exemple, Parrila *et al.* (2004) ont obtenu un R^2 de 0.425 dans le cadre de leur recherche évaluant le rôle prédictif de certaines variables (décodage, MCT verbale, dénomination et conscience phonologique) en première année sur la compréhension en 3e année. Tunmer et Chapman (2012) ont évalué le rôle de l'âge, la compréhension linguistique, le décodage et le vocabulaire dans les performances en compréhension de la lecture. Leur modèle de régression hiérarchique explique près de 80 % de la variance en compréhension, avec un R^2 de 0.797.

4.1.3.1 Décodage

La mesure finale de décodage, utilisée dans notre modèle de régression linéaire comme variable dépendante ou prédite, correspond à la moyenne des résultats en score Z pour les deux épreuves de décodage, soit la lecture de mots en une minute et la lecture de non-mots en une minute. Le tableau des corrélations montre que le décodage est peu lié aux différentes épreuves. Néanmoins, une analyse de régression est souhaitable, afin que nous soyons en mesure de comparer les différents modèles. Rappelons que nos modèles de régression impliquant le temps 1 et le temps 2 ne se sont pas avérés significatifs. Les variables prédictives dans le tableau 4.6 correspondent donc aux épreuves du temps 3.

Tableau 4.6 – Résultats pour les analyses de régression pour le décodage au temps 3

| T3 et Décodage | Coefficients non standardisés | | Coefficients standardisés | | Sig. |
|---------------------------------------|-------------------------------|-----------------|---------------------------|--------|--------|
| | A | Erreur standard | Bêta | t | |
| (Constante) | -2,284 | 0,933 | | -2,449 | 0,019 |
| Répétition de mots rares | -0,826 | 1,055 | -0,17 | -0,782 | 0,439 |
| Répétition de chiffres | 0,547 | 0,934 | 0,133 | 0,586 | 0,561 |
| Répétition de rythmes | 1,891 | 0,958 | 0,336 | 1,973 | 0,055 |
| Topologie et arithmétique | -2,398 | 0,729 | -0,629 | -3,29 | 0,002* |
| Vocabulaire (expression et réception) | 3,473 | 1,252 | 0,488 | 2,774 | 0,008* |

Le modèle permet d'expliquer 27% de la variance en décodage ($R = 0,520$; $R^2 = 0,271$; R^2 ajusté = 0,18; $F(5,45) = 2,972$; $p < 0,023$). Dans ce modèle, ce sont les contributions des épreuves topologie et arithmétique ($\beta = -0,629$; $p < 0,002$) et de vocabulaire (expression et réception) ($\beta = 0,488$; $p < 0,008$) qui sont significatives. On dégage de l'analyse un patron où l'épreuve de topologie et arithmétique influence négativement les résultats en décodage et où l'épreuve de vocabulaire exerce sur la variable dépendante une influence positive. Il s'ensuit qu'un individu ayant une bonne performance à l'épreuve de topologie et arithmétique présentera des lacunes en décodage. Les épreuves de répétition de mots rares, répétition de chiffres et répétition de rythmes n'ont pas d'influence significative sur la variable finale.

4.1.3.2 Compréhension

La mesure finale de compréhension en lecture, utilisée ici comme variable prédite, correspond au score pour l'épreuve de compréhension de textes (9 textes). Les tableaux des corrélations ont montré que la mesure de la compréhension était faiblement corrélée avec certaines des différentes épreuves des temps 1 et 2 (voir les

tableaux 4.2 et 4.3). Le tableau 4.4 a mis en lumière des corrélations légèrement plus fortes entre les épreuves de vocabulaire au temps 3 (expression et réception) et la compréhension en temps final. Nous procéderons à des analyses de régression afin d'être en mesure de comparer les modèles des temps 1, 2 et 3.

Tableau 4.7 – Résultats pour les analyses de régression pour la compréhension au temps 1

| T1 et Compréhension | Coefficients non standardisés | | Coefficients standardisés | | |
|---------------------------------------|-------------------------------|-----------------|---------------------------|--------|--------|
| | A | Erreur standard | Bêta | t | Sig. |
| (Constante) | 0,305 | 0,117 | | 2,617 | 0,013 |
| Répétition de mots rares | -0,152 | 0,1 | -0,244 | -1,515 | 0,138 |
| Répétition de chiffres | -0,034 | 0,105 | -0,053 | -0,322 | 0,749 |
| Topologie et arithmétique | -0,076 | 0,103 | -0,109 | -0,737 | 0,466 |
| Répétition de phrases | 0,396 | 0,125 | 0,591 | 3,166 | 0,003* |
| Vocabulaire (expression et réception) | 0,402 | 0,195 | 0,313 | 2,063 | 0,046* |

Le modèle du temps 1, dont les résultats figurent dans le tableau 4.7, a permis d'expliquer 38% de la variance dans les résultats en compréhension, s'est avéré significatif ($R = 0,615$; $R^2 = 0,378$; R^2 ajusté = $0,298$; $F(5,44) = 4,737$; $p < 0,002$). Les épreuves de répétition de phrases ($\beta = 0,591$; $p < 0,003$) et de vocabulaire ($\beta = 0,313$; $p < 0,046$) exercent une influence positive et significative sur la variance en compréhension à l'intérieur de ce modèle; la première possède par ailleurs un bêta presque deux fois plus fort que la deuxième. Les épreuves de répétition de mots rares, répétition de chiffres et de topologie et arithmétique n'ont pas exercé une influence significative sur la variable finale.

L'autre modèle qui a été soumis à l'analyse avait comme variables prédictives les épreuves du temps 2 et comme variable prédite, la compréhension. Le modèle s'est avéré significatif ($R = 0,538$; $R^2 = 0,29$; R^2 ajusté = $0,194$; $F(5,42) = 3,02$; $p < 0,022$) et

contribue à expliquer 29% de la variance dans les performances en compréhension. Le modèle du temps 2 est donc moins à même d'expliquer la variance dans les résultats en compréhension en temps final que le modèle du temps 1. Les détails sont présentés dans le tableau 4.8.

Tableau 4.8 – Résultats pour les analyses de régression pour la compréhension au temps 2

| T2 et Compréhension | Coefficients non standardisés | | Coefficients standardisés | | |
|---------------------------------------|-------------------------------|-----------------|---------------------------|--------|--------|
| | A | Erreur standard | Bêta | t | Sig. |
| (Constante) | 0,408 | 0,104 | | 3,925 | 0 |
| Répétition de mots rares | -0,062 | 0,194 | -0,058 | -0,318 | 0,752 |
| Répétition de chiffres | 0,014 | 0,118 | 0,022 | 0,119 | 0,906 |
| Topologie et arithmétique | -0,192 | 0,088 | -0,371 | -2,18 | 0,036 |
| Répétition de phrases | 0,091 | 0,193 | 0,103 | 0,471 | 0,641 |
| Vocabulaire (expression et réception) | 0,594 | 0,196 | 0,607 | 3,027 | 0,004* |

L'épreuve de répétition de phrases en temps 2 n'est pas parvenue à influencer significativement les performances en compréhension. Il en est de même pour les épreuves de répétition de mots rares et de répétition de chiffres. L'épreuve de vocabulaire ($\beta=0,607$; $p<0,004$) exerce une influence positive et significative sur la variance en compréhension à l'intérieur de ce modèle. L'épreuve de topologie et arithmétique ($\beta=-0,371$; $p<0,036$) a eu une influence négative et significative sur la variable finale.

Dans le tableau 4.9 figurent les détails des résultats pour le dernier modèle de régression que nous considérerons dans la présente section. Il avait comme variables prédictives les épreuves du temps 3 et comme variable prédite, la compréhension.

Tableau 4.9 – Résultats pour les analyses de régression pour la compréhension au temps 3

| | Coefficients non standardisés | | Coefficients standardisés | | |
|---------------------------------------|-------------------------------|-----------------|---------------------------|--------|---------|
| | A | Erreur standard | Bêta | t | Sig. |
| T3 et Compréhension | | | | | |
| (Constante) | 0,125 | 0,134 | | 0,937 | 0,355 |
| Répétition de mots rares | 0,156 | 0,151 | 0,183 | 1,033 | 0,308 |
| Répétition de chiffres | -0,001 | 0,127 | -0,002 | -0,011 | 0,991 |
| Répétition de rythmes | -0,106 | 0,127 | -0,109 | -0,828 | 0,412 |
| Topologie et arithmétique | -0,074 | 0,1 | -0,119 | -0,744 | 0,461 |
| Vocabulaire (expression et réception) | 0,841 | 0,18 | 0,696 | 4,667 | 0,0001* |

Ce modèle est parvenu à expliquer 50% de la variance dans les résultats en compréhension ($R= 0,703$; $R^2= 0,495$; R^2 ajusté= $0,432$; $F(5,45)=7,838$; $p<0,0001$). Seule l'épreuve de vocabulaire a présenté une influence significative sur la compréhension au temps 3 ($\beta= 0,696$; $p<0,0001$). Ni l'épreuve de répétition de mots rares, ni l'épreuve de répétition de chiffres, ni l'épreuve de répétition de rythmes ou l'épreuve de topologie et arithmétique du temps 3 ne sont parvenues à jouer un rôle significatif dans l'explication de la variance de la variable compréhension en temps final.

4.1.3.3 Synthèse des résultats obtenus pour les analyses effectuées à partir des variables mesurées

Les analyses contenues dans cette section de notre travail de recherche nous ont permis de constater que certaines variables n'exercent jamais d'influence significative sur les performances en lecture. Il s'agit des épreuves de répétition de mots rares, de répétition de chiffres et de répétition de rythmes. Ces épreuves correspondent à trois de nos quatre mesures de la mémoire à court terme, temps 1, 2 et 3 confondus. De plus, soulignons que globalement, les modèles du temps 3 fonctionnent mieux : ils sont significatifs et expliquent une meilleure partie de la variance dans les résultats en

lecture, en décodage et en compréhension. Cet état de choses suggère que les données recueillies en temps 3 sont de meilleurs prédicteurs de la performance en temps final, ce qui s'explique probablement en partie par la proximité temporelle des deux passations. Au temps 1, l'épreuve de répétition de phrases s'impose comme le meilleur prédicteur des performances en lecture et en compréhension, conformément à ce qu'avaient trouvé les études de Butler *et al.* (1985) et de Catts, Fey et Zhang (2002). Cette épreuve constitue donc un bon outil de dépistage des problèmes ultérieurs en compréhension de la lecture.

4.2 Analyses appliquées aux habiletés latentes

Les analyses contenues dans cette section de notre mémoire nous permettront de répondre à notre deuxième question de recherche, dont l'objectif est d'expliquer le rapport entre nos deux habiletés latentes (l'une cognitive – la MCT – et l'autre langagière – le vocabulaire) et les performances en compréhension de la lecture et en décodage. Pour ce faire, nous devons au préalable isoler ces deux habiletés transversales à partir des résultats à chacune des épreuves de base en T1, T2 et T3.

4.2.1 ACP permettant la construction des variables latentes

Cette section est dédiée aux analyses factorielles en composante principale (ACP) auxquelles nous avons procédé. Ce type d'analyses permet de voir comment les données s'organisent, en construisant plusieurs axes, ou composantes, qui sont orthogonaux, et, de ce fait, non-corrélés. En effet, avec les ACP, les données sont soumises à un traitement où elles doivent s'ajuster au mieux à des composantes qui sont indépendantes entre elles. Ces analyses nous ont donc permis d'abord de répondre à nos deux sous-questions de recherche, qui visent à déterminer si les épreuves parviennent à construire des habiletés latentes qui nous permettent de mieux

comprendre les forces à l'œuvre dans les performances en lecture. En effet, les ACP permettent de voir comment chaque variable est bien représentée par l'un ou l'autre des axes, laissant à la chercheuse ou au chercheur le soin de déterminer ce que représente l'habileté mise en évidence.

Nous avons procédé aux ACP avec méthode de rotation Varimax pour les temps 1 (figure 4.1), 2 (figure 4.2) et 3 (figure 4.3) afin de voir comment se structuraient les variables. La façon dont les épreuves sont agencées nous donne les outils pour parvenir à une interprétation des axes. Les graphiques suivants, qui représentent les ACP pour les temps 1, 2 et 3, permettent de constater des tendances claires. D'abord, les épreuves de vocabulaire (expression et réception) et topologie et arithmétique semblent être bien représentées par le premier axe. Les épreuves de répétition (répétition de phrases, de mots rares, de chiffres et de rythmes dans le cas du temps 3), quant à elles, tendent à construire le deuxième axe. Pour tous les temps, c'est l'épreuve de répétition de phrases qui semble être la mesure la plus composite, étant plutôt bien représentée par les deux axes.

Tableau 4.10 – Statistiques relatives aux composantes créées par ACP aux temps 1, 2 et 3

| Temps | Composante | Valeurs propres initiales | |
|---------|------------|---------------------------|------------------|
| | | Total | % de la variance |
| Temps 1 | 1 | 2,922 | 48,698 |
| | 2 | 1,245 | 20,746 |
| Temps 2 | 1 | 3,123 | 52,051 |
| | 2 | 1,220 | 20,332 |
| Temps 3 | 1 | 3,249 | 54,151 |
| | 2 | 1,241 | 20,676 |

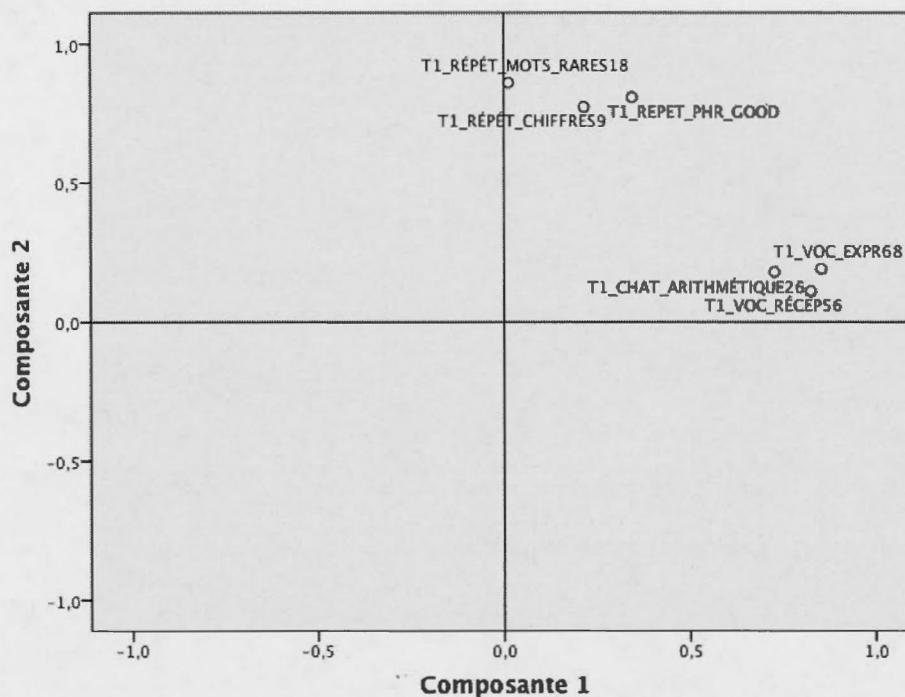


Figure 4.1 – Diagramme de composantes (temps 1)

Les valeurs propres associées à chaque axe, qui sont reportées dans la figure 4.1, nous permettent de déterminer combien de facteurs sont à prendre en considération dans notre analyse. Ces valeurs représentent le nombre d'éléments qui sont expliqués par le facteur. Selon cette logique, nous ne considérons que les facteurs expliquant au moins un élément, soit ceux dont la valeur propre est au-dessus de 1 (Lebart, Morineau et Piron, 1995, p. 54). Les trois analyses ont permis, chaque fois, d'isoler deux facteurs expliquant la variance observée. À tous les temps, le premier facteur possède une valeur propre d'environ 3, alors que celle du deuxième facteur se chiffre à 1,2. Le premier facteur, qui est le plus important, explique tour à tour 48,7% (temps 1), 52,1% (temps 2) et 54,2% (temps 3) de la variance. Ce facteur est représenté aux temps 1 et 2 par les mêmes trois épreuves, soit les épreuves de vocabulaire (expression et réception) et l'épreuve de topologie et arithmétique, qui vise à évaluer

la présence d'un vocabulaire réceptif bien ciblé. Au temps 3, le facteur principal est construit sensiblement par les trois mêmes épreuves, soit les épreuves de topologie et de vocabulaire (expression et réception). Le deuxième facteur, quant à lui, explique 20,7 % de la variance au temps 1, 20,3 % au temps 2 et 20,7 % au temps 3.

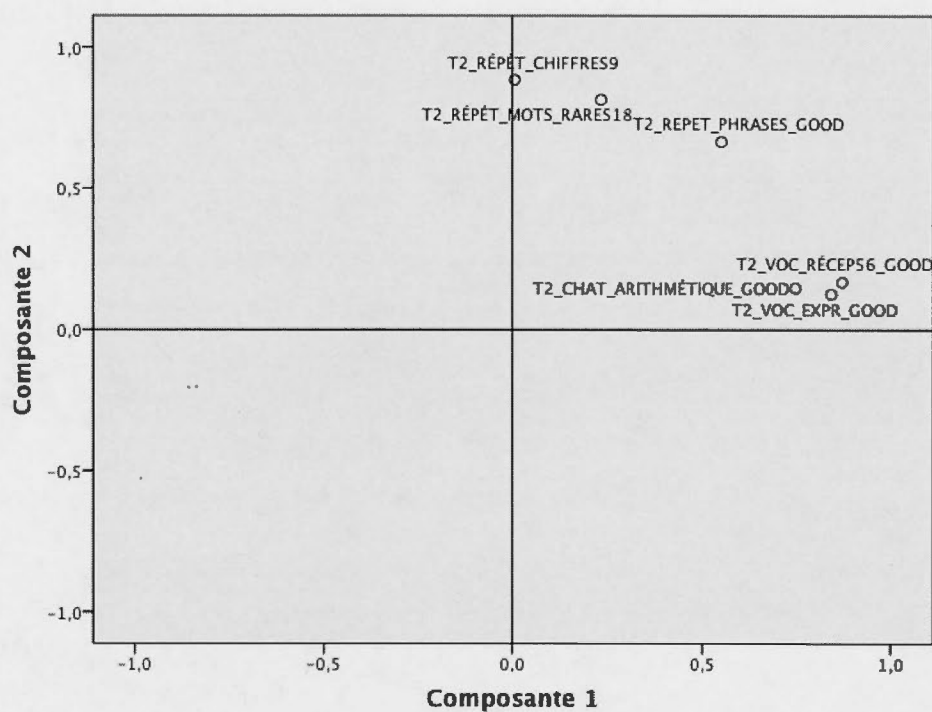


Figure 4.2 – Diagramme de composantes (temps 2)

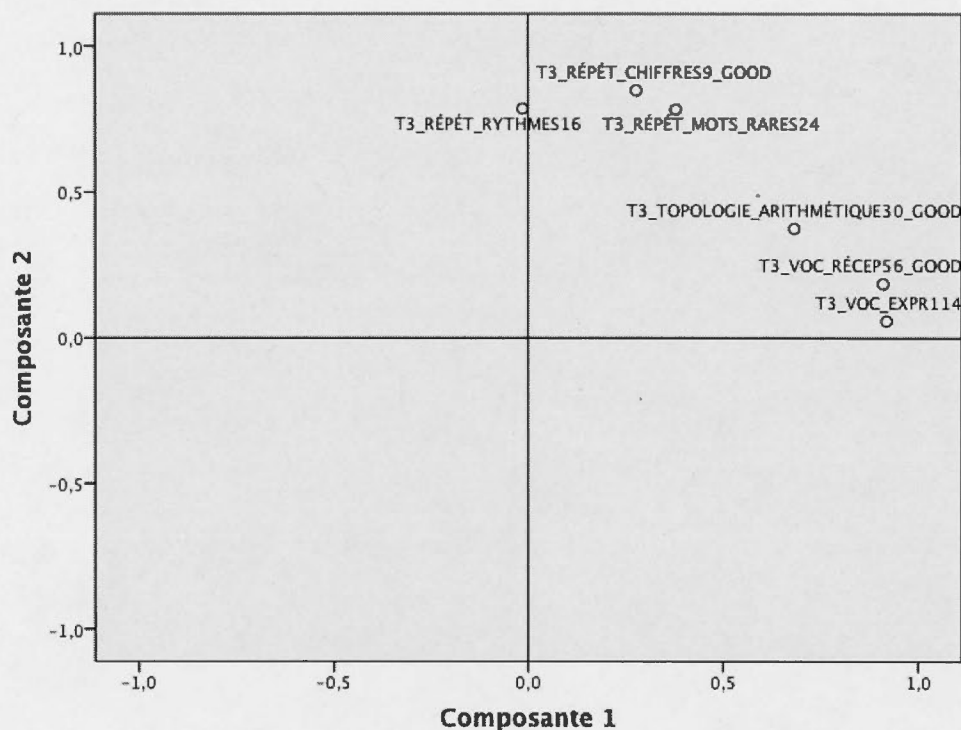


Figure 4.3 – Diagramme de composantes (temps 3)

La similitude des résultats observés pour les trois temps est frappante. Les ACP effectuées nous permettent de dégager un profil cohérent : elles mettent en évidence 2 facteurs qui organisent les 6 variables sur chacun des temps. Les figures 4.1, 4.2 et 4.3 présentent des tendances claires. Les deux facteurs qui émergent de l'analyse, représentés à chaque temps par les mêmes éléments essentiellement, semblent se référer à deux grandes habiletés, orthogonales ou non : d'une part, le vocabulaire, et d'autre part, la MCT ou, plus prudemment, l'habileté à répéter des items. Cette exploration préliminaire de nos variables nous a permis de répondre à nos sous-questions de recherche. Nous pouvons en effet constater empiriquement que les épreuves parviennent à construire deux habiletés latentes homogènes qui

correspondent au vocabulaire et à la MCT. De plus, ces analyses légitiment les regroupements qui nous permettront de répondre à notre deuxième question de recherche, à savoir comment ces habiletés latentes, mesurées en maternelle et en 1^{re} année, contribuent à expliquer les performances en décodage et en compréhension de la lecture à la fin de la 3^e et de la 4^e année du primaire.

4.2.2 Statistiques descriptives des variables latentes construites par ACP

Les analyses descriptives permettent de vérifier que chaque variable que les ACP nous ont permis de construire une distribution dite « normale ». Rappelons qu'une distribution normale se caractérise par des valeurs d'asymétrie et de l'aplatissement entre -1,5 et 1,5 (Carricano *et al.*, 2010). Le tableau 4.11 atteste le fait que les variables utilisées respectent ces pré-requis. Par ailleurs, les ACP fournissent, pour chaque élève, une position sur les axes. Ainsi, sur un même plan, il est possible de projeter chacun des élèves, grâce aux coordonnées que nous possédons. Ces coordonnées correspondent à la mesure de l'élève sur l'axe vocabulaire (axe 1) et sur l'axe mémoire (axe 2). Ces valeurs peuvent notamment servir à analyser de façon personnalisée les lacunes et les forces d'un ou d'une élève en particulier par rapport aux deux axes.

4.2.3 Analyses de corrélation des variables latentes construites par ACP

Cette section fait état des liens de corrélations qui unissent les habiletés cognitive et langagière construites et les performances en compréhension de la lecture et en décodage. Le tableau qui suit regroupe les valeurs des coefficients de corrélation (r) de Pearson de l'ensemble des variables qui nous intéressent. Ici, ce sont les variables latentes de la MCT et du vocabulaire aux temps 1, 2 et 3 qui sont croisées avec les variables finales en décodage et en compréhension.

Tableau 4.11 – Statistiques descriptives des variables latentes à l'étude retenues pour l'analyse

| Variabes latentes | Min | Max | Asymétrie | Aplatissement | N |
|-------------------|----------|---------|-----------|---------------|----|
| T1 Vocabulaire | -2,10063 | 2,23072 | 0,451 | -0,172 | 50 |
| T1 MCT | -1,82007 | 2,32151 | 0,421 | -0,413 | 50 |
| T2 Vocabulaire | -2,40493 | 1,96466 | 0,162 | -0,222 | 48 |
| T2 MCT | -2,45117 | 1,77123 | -0,227 | 0,001 | 48 |
| T3 Vocabulaire | -2,24182 | 1,72086 | -0,166 | -0,607 | 53 |
| T3 MCT | -2,16274 | 2,50079 | 0,188 | 0,072 | 53 |

Note : La manipulation statistique à laquelle nous avons soumis nos données fait en sorte que les deux facteurs possèdent une moyenne de 0 et un écart-type de 1. C'est pour cette raison que nous ne faisons pas état de ces valeurs dans le tableau.

Tableau 4.12 – Corrélations entre toutes les variables (Temps 1)

| Variabes | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
|------------------|----------|---------|----------|---------|---------|---------|---------|---------|----|
| | <i>r</i> | 1 | | | | | | | |
| 1. T1 Voc | <i>n</i> | 50 | | | | | | | |
| | <i>r</i> | 0 | 1 | | | | | | |
| 2. T1 MCT | <i>n</i> | 50 | 50 | | | | | | |
| | <i>r</i> | 0,748** | 0,303* | 1 | | | | | |
| 3. T2 Voc | <i>n</i> | 48 | 48 | 48 | | | | | |
| | <i>r</i> | 0,07 | -0,477** | 0 | 1 | | | | |
| 4. T2 MCT | <i>n</i> | 48 | 48 | 48 | 48 | | | | |
| | <i>r</i> | 0,738** | 0,175 | 0,823** | 0,284 | 1 | | | |
| 5. T3 Voc | <i>n</i> | 43 | 43 | 41 | 41 | 53 | | | |
| | <i>r</i> | 0,128 | -0,511** | -0,161 | 0,721** | 0 | 1 | | |
| 6. T3 MCT | <i>n</i> | 43 | 43 | 41 | 41 | 53 | 53 | | |
| | <i>r</i> | 0,094 | 0,047 | -0,046 | -0,016 | 0,094 | -0,054 | 1 | |
| 7. Décodage | <i>n</i> | 48 | 48 | 46 | 46 | 48 | 48 | 70 | |
| | <i>r</i> | 0,281 | 0,076 | 0,165 | -0,102 | 0,464** | -0,314* | 0,359** | 1 |
| 8. Compréhension | <i>n</i> | 48 | 48 | 46 | 46 | 48 | 48 | 70 | 70 |

** La corrélation est significative au niveau 0.01 (bilatéral).

* La corrélation est significative au niveau 0.05 (bilatéral).

Nous commencerons par une description des liens de corrélation qu'entretiennent les deux variables finales avec le reste des variables. D'abord, remarquons les corrélations non-significatives que le décodage entretient avec les différentes habiletés latentes des 3 temps initiaux. Par ailleurs, une corrélation moyenne unit le décodage et la compréhension en lecture (r de 0,359; $p < 0,01$). La variable représentant les performances en compréhension en lecture est corrélée avec la variable du vocabulaire au temps 3 (r de 0,464; $p < 0,01$). Le fait que la compréhension en lecture soit corrélée négativement avec la MCT au temps 3 est une observation inédite, à notre connaissance, dans la littérature. Daneman et Carpenter (1980) font état de corrélations faibles voire nulles pour les mêmes variables. La variable latente de la compréhension n'entretient pas un tel patron de corrélations avec les variables des temps 1 et 2 : les corrélations sont faibles et non-significatives. Nous nous pencherons sur cet état de choses dans la discussion.

Nous nous attarderons maintenant à la façon dont les différentes variables des temps 1, 2 et 3 interagissent. Observons d'entrée de jeu que celles qui se réfèrent à l'habileté générale en vocabulaire aux temps 1, 2 et 3 sont fortement corrélées (T1 et T2 : r de 0,748; $p < 0,01$ | T2 et T3 : r de 0,823; $p < 0,01$ | T1 et T3 : r de 0,738; $p < 0,01$). La variable de MCT du temps 1 est corrélée négativement avec les variables analogues des temps 2 et 3 (T1 et T2 : r de -0,477; $p < 0,01$ | T1 et T3 : r de -0,511; $p < 0,01$). Pourtant, la corrélation de la variable MCT des temps 2 et 3 sont corrélées positivement et fortement (r de 0,721; $p < 0,01$). Cette petite incohérence des axes Mémoire sera aussi traitée dans la discussion.

4.2.4 Analyses de régression et analyses en composantes principales (ACP)

Les analyses en composantes principales présentées dans la section 4.2.1 ont permis de créer deux axes ou facteurs orthogonaux. La revue de la littérature (chapitre II), les

discussions sur la nature des épreuves et sur ce qu'il faut mobiliser comme ressources pour les réussir et les résultats des analyses statistiques nous conduisent à considérer l'axe 1 comme une représentation empirique du vocabulaire et l'axe 2 comme une représentation empirique de la MCT. La présente section de l'étude poursuit l'objectif de nous donner les outils pour répondre à notre dernière question de recherche, soit de déterminer de quelle façon la mémoire à court terme et le vocabulaire aux temps 1, 2 et 3 contribuent à expliquer les performances en décodage et en compréhension au temps final. Pour ce faire, nous avons procédé à des analyses de régression avec les deux axes créés par les analyses en composantes principales comme variable indépendantes ou prédictives (voir tableau 4.13). Les ACP nous ont permis de créer des variables latentes, construites avec au moins 3 variables manifestes – les épreuves. Ces analyses synthétisent l'information provenant des variables mesurées et construisent des composantes plus pures. Rappelons que les variables latentes en question constituent nos mesures des habiletés que nous souhaitons mettre en lien avec les composantes de la lecture : la MCT et le vocabulaire. Les ACP nous ont fourni les coordonnées de chacun des élèves pour les deux axes, aux temps 1, 2 et 3. Ces valeurs correspondent concrètement à la position graphique des élèves par rapport à ces deux axes, sur les trois temps. Ce sont ces données que nous considérerons comme des variables prédictives dans les analyses de régression subséquentes.

Encore une fois, les données du temps 2 ne génèrent pas de modèles significatifs. Aucun modèle n'incluant le décodage ne s'est avéré significatif. Les deux modèles dont nous ferons état des résultats sont ceux du temps 1 et du temps 3 avec comme variable dépendante la compréhension. Soulignons à nouveau que nous avons évalué le pouvoir explicatif des deux modèles après suppression des résidus.

Tableau 4.13 – Six modèles de régression testés

| Temps | Variable prédite | Modèle 1 | | | Modèle 2 (après suppression des résidus) | | | | | |
|-------|--------------------------|----------|----------------|----|--|----------------|----|----------------|----------|------|
| | | R | R ² | n | R | R ² | n | F | P | Gain |
| T1 | Com- préhen- -sion | 0,29 | 0,084 | 48 | 0,44 | 0,194 | 46 | F(2,45)=5,166 | 0,01* | 11% |
| T2 | | 0,194 | 0,038 | 46 | 0,299 | 0,09 | 45 | F(2,43)=2,02 | 0,146 | 5% |
| T3 | | 0,549 | 0,301 | 45 | 0,698 | 0,488 | 42 | F(2,44)=19,978 | 0,0001** | 19% |
| T1 | Déco- dage | 0,105 | 0,011 | 48 | 0,218 | 0,048 | 47 | F(2,45)=1,075 | 0,35 | 4% |
| T2 | | 0,049 | 0,002 | 46 | 0,019 | 0 | 45 | F(2,44)=0,008 | 0,992 | 0% |
| T3 | | 0,107 | 0,011 | 45 | 0,107 | 0,011 | 45 | F(2,47)=0,259 | 0,773 | 0% |

Au temps 1, le modèle (tableau 4.14) parvient à expliquer 19 % de la variance dans les résultats en compréhension ($R = 0,440$; $R^2 = 0,194$; R^2 ajusté = $0,156$; $F(2,45) = 5,166$; $p < 0,01$). La variable construite en majeure partie par des épreuves de vocabulaire, grâce aux analyses en composantes principales, a présenté une influence significative sur la compréhension au temps 1 ($\beta = 0,424$; $p < 0,003$). La variable construite, que nous avons assimilée à une mesure de la MCT, ne joue pas un rôle significatif dans l'explication de la variance de la variable compréhension au temps 1 ($\beta = 0,12$). La régression impliquant les variables construites à partir des mesures du temps 3 et la compréhension en temps final comme variable dépendante a généré un modèle significatif ($R = 0,698$; $R^2 = 0,488$; R^2 ajusté = $0,463$; $F(2,44) = 19,978$; $p < 0,00001$) qui permet d'expliquer 49 % de la variance de la variable finale (tableau 4.15). L'axe vocabulaire contribue de façon significative à la construction de la compréhension ($\beta = 0,594$; $p < 0,00001$), mais l'axe MCT présente une contribution négative et significative à celle-ci ($\beta = -0,293$; $p < 0,012$).

Tableau 4.14 Résultats pour les analyses de régression pour la compréhension au temps 1

| | T1 et Compréhension | Coefficients non standardisés | | Coefficients standardisés | t | Sig. |
|----|------------------------|----------------------------------|-----------------|------------------------------|-------|--------|
| | | A | Erreur standard | Bêta | | |
| | (Constante) | 0,221 | 0,123 | | 1,799 | 0,079 |
| V1 | Axe Vocabulaire | 0,39 | 0,126 | 0,424 | 3,098 | 0,003* |
| M1 | Axe MCT | 0,105 | 0,119 | 0,12 | 0,879 | 0,384 |

Tableau 4.15 – Résultats pour les analyses de régression pour la compréhension au temps 3

| | T3 et Compréhension | Coefficients non standardisés | | Coefficients standardisés | t | Sig. |
|----|------------------------|----------------------------------|-----------------|------------------------------|--------|---------|
| | | A | Erreur standard | Bêta | | |
| | (Constante) | 0,149 | 0,107 | | 1,393 | 0,171 |
| V3 | Axe Vocabulaire | 0,6 | 0,113 | 0,594 | 5,324 | 0,0001* |
| M3 | Axe MCT | -0,285 | 0,109 | -0,293 | -2,627 | 0,012* |

Le vocabulaire s'impose donc, dans les deux modèles observés, comme une habileté contribuant significativement à expliquer les performances en compréhension de la lecture. Nous ne pouvons commenter à ce moment-ci les résultats pour le décodage, puisque les modèles avec les variables latentes des temps 1, 2 et 3 et le décodage n'étaient pas significatifs. La figure 4.4 permet de visualiser les résultats que nous avons obtenus concernant notre deuxième question de recherche, de nature explicative. Nous examinerons minutieusement ces résultats dans le chapitre suivant.

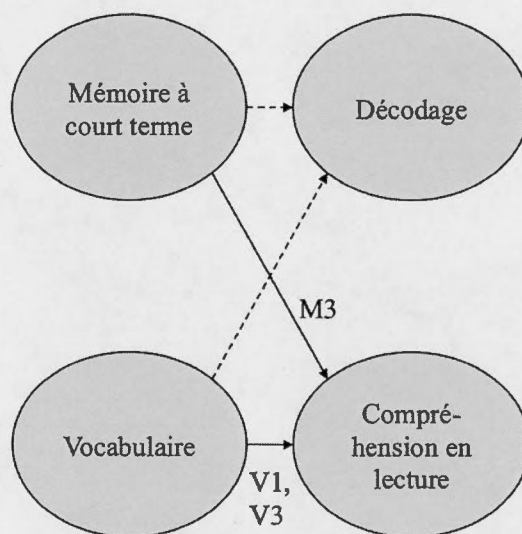


Figure 4.4 – Modélisation des résultats pour les analyses de régression effectuées avec les variables latentes

Note 1 : Les codes V1, V3 et M3 se réfèrent à ceux dans les tableaux des analyses de régression 4.14 et 4.15.

Note 2 : Les traits pointillés correspondent à des contributions présentes dans le modèle mais qui ne se sont pas avérées significatives.

CHAPITRE V

DISCUSSION

C'est dans cette section que nous répondrons à nos questions de recherche, en faisant une analyse de nos résultats présentés dans le chapitre IV, au regard des modèles théoriques et écrits scientifiques recensés au chapitre II. Nous déterminerons d'abord dans quelle mesure chacune des épreuves soumises aux participants de la maternelle à la 1^{re} année prédisent les habiletés en lecture en 3^e et 4^e année, ce qui répondra à notre première question de recherche. Ensuite, nous tenterons d'expliquer comment le vocabulaire et la mémoire à court terme contribuent à expliquer les performances en lecture en 3^e et en 4^e année. Ceci répondra à notre deuxième question de recherche. Nous compléterons ce chapitre par une revue critique de la méthodologie utilisée dans le cadre de ce mémoire et nous questionnerons sur la portée et les limites de notre recherche.

5.1 Réponse à la première question de recherche

Notre première question de recherche visait à déterminer quelles étaient les épreuves, parmi celles des temps 1, 2 et 3, nous permettant de mieux prédire les performances en lecture. Pour répondre à cette question, nous avons réalisé des analyses de régression avec les épreuves de chaque temps (1, 2 et 3), avec comme variables prédites la compréhension en lecture et le décodage.

Le tableau 5.1 synthétise la réponse à notre première question de recherche. Il indique les épreuves qui se sont révélées être des dépisteurs efficaces des performances ultérieures en compréhension en lecture.

Tableau 5.1 – Synthèse des résultats sur les épreuves des T1, T2 et T3 permettant de prédire les performances en décodage et en compréhension au TF

| Variable prédite (TF) | Meilleures épreuves (variables prédictives) | | Bêta |
|-----------------------|---|---------------------------------------|--------|
| Compréhension | T1 | Répétition de phrases | 0,591 |
| | | Vocabulaire (expression et réception) | 0,313 |
| | T2 | Vocabulaire (expression et réception) | 0,607 |
| | | Topologie et arithmétique | -0,371 |
| T3 | Vocabulaire (expression et réception) | 0,696 | |
| | T3 | Topologie et arithmétique | -0,629 |
| Décodage | | Vocabulaire (expression et réception) | 0,488 |

5.1.1 Épreuves permettant de prédire les performances en compréhension

Penchons-nous d'abord sur les épreuves susceptibles de prédire le mieux les performances finales en compréhension de la lecture. Notons que les compétences en compréhension de la lecture regroupent un ensemble de processus de haut niveau qui sont liés à la réussite d'ordres scolaire, personnel et social (Desrochers et Berger, 2011). La prédiction de telles habiletés dans une perspective longitudinale à l'aide d'épreuves est donc d'une importance capitale pour la communauté que nous avons étudiée. Soulignons qu'un certain nombre d'épreuves ne sont pas parvenues, du moins, d'un point de vue longitudinal, à expliquer les performances en lecture. Il s'agit des épreuves de répétition de mots rares, de répétition de chiffres et de répétition de rythmes.

Les analyses de corrélation avaient montré des liens faibles à moyens entre certaines épreuves – celles de vocabulaire (expression et réception) aux temps 1, 2 et 3, de

répétition de phrases (temps 1) et de topologie et arithmétique au temps 3 – et la compréhension en lecture. Les analyses de régression ont mis en relief le pouvoir prédictif de ces épreuves.

L'épreuve de répétition de phrases mesurée à 5 ans (T1) est un prédicteur fiable des performances en compréhension de la lecture à 9 et 10 ans (TF) ($\beta = 0,591$). Ces résultats concordent avec ceux de Butler *et al.* (1985) et de Catts *et al.* (2001). Cette épreuve est considérée comme l'une des plus efficaces dans la prédiction des performances en lecture, particulièrement pendant les premières années de scolarisation (Butler *et al.*, 1985). Catts *et al.* (2001) ont pour leur part observé que l'épreuve de répétition de phrases, évaluée en maternelle, contribuait de façon unique à expliquer la variance dans les résultats en lecture en 2^e année. Mentionnons que Catts *et al.* (2001) ont considéré cette épreuve comme une mesure de l'étendue du vocabulaire. Willis et Gathercole (2001) ont, quant à elles, confirmé que la capacité à répéter des phrases est fortement corrélée avec la mesure de la boucle phonologique. Il s'agit d'un véritable problème théorique. Les unités que l'épreuve met en œuvre sont de natures diverses (syntaxe, mémoire, phonologie, sémantique) : l'épreuve mobilise certainement plus d'une habileté. Les auteurs de la batterie NÉEL (Chevrie-Muller et Plaza, 2001) soulignaient d'ailleurs d'emblée la complexité et la variété des habiletés mises en œuvre lors de la passation de cette épreuve. Par ailleurs, l'épreuve mobilise des habiletés de haut niveau : en mesurant tant des processus cognitifs, que syntaxiques et sémantiques, l'épreuve a effectivement toutes les chances de prédire avec force les performances en compréhension.

Les autres épreuves qui ont contribué à expliquer les performances ultérieures en compréhension sont toutes des épreuves mesurant l'étendue du vocabulaire. Nous aborderons le lien privilégié entre vocabulaire et compréhension en lecture dans la section 5.2.1.

5.1.2 Épreuves permettant de prédire les performances en décodage

Seul un modèle de régression a permis d'expliquer significativement les performances en décodage. Il s'agit de celui du temps 3. Le modèle permet d'expliquer 27% de la variance. Les résultats en décodage n'étaient pas ou peu corrélés avec ceux des épreuves des temps 1, 2 et 3. Les deux épreuves de vocabulaire incluses dans le modèle ont eu une influence significative sur le décodage, l'une négativement – topologie et arithmétique : $\beta = -0,629$ – et l'autre positivement – vocabulaire expressif et réceptif : $\beta = 0,488$. Les natures différentes de ces épreuves sont probablement la cause de ce résultat contrasté. Dans les différents sous-tests qui construisent l'épreuve de topologie et arithmétique, la réponse est surtout de nature réceptive et implique de concevoir l'espace : l'enfant est amené à identifier les éléments demandés du doigt, à manipuler des figurines d'animaux et à compter des éléments. L'épreuve demande à ce que l'enfant se situe par rapport à la situation qu'on l'oblige à analyser et qu'il connaisse surtout des mots-liens (« le même », « entre », « l'un derrière l'autre », « face à face », etc.) Pour réussir, il doit créer une image mentale de la situation. À l'opposé, l'épreuve de décodage mesure l'automatisation d'un processus basé sur le transcodage de petites unités; l'accès au sens est relégué à une étape ultérieure dans le processus de lecture. L'influence négative de l'épreuve de topologie et arithmétique sur le décodage pourrait s'expliquer par le fait que le participant doit s'extraire de l'unité pour aller vers le global. Dans les autres épreuves de vocabulaire, les réponses sont de nature réceptive et productive et abordent différents thèmes. Le champ lexical est élargi : l'enfant doit nommer par lui-même et indiquer du doigt des éléments de la vie quotidienne, comme des objets dans la salle de bains ou la cuisine, et des éléments qui font l'objet d'un enseignement explicite à l'école, comme les couleurs, les formes et les parties du corps. Notre revue de la littérature sur le vocabulaire nous donne à penser que la

deuxième épreuve, celle de vocabulaire expressif et réceptif, est probablement plus à même de mesurer adéquatement les habiletés en étendue du vocabulaire. En effet, cette épreuve est plus diversifiée que les autres sur le plan de la modalité des réponses et des thèmes traités. De plus, l'épreuve de vocabulaire expressif bénéficie d'une méthode de pointage graduée (0, 1 ou 2 points) qui permet d'affiner davantage les résultats.

Nous constatons que l'épreuve de vocabulaire (expression et réception) constitue un bon prédicteur des performances en décodage, à l'intérieur d'un modèle incluant des épreuves mobilisant d'autres types d'habiletés. Ce lien entre une mesure du vocabulaire et le décodage n'est pas inédit, mais reste peu documenté. Beck et McKeown (1991) (cité dans Aarnoutse *et al.*, 2001, p. 63) ont repéré un effet de la taille du vocabulaire sur les performances en décodage et en compréhension en lecture. Ricketts *et al.* (2007) avaient, pour leur part, observé que le vocabulaire ne constituait pas un prédicteur d'habiletés de bas niveau en lecture, nommément le décodage.

5.2 Réponse à la deuxième question de recherche

Notre deuxième question de recherche avait pour but de vérifier la contribution de la MCT et du vocabulaire aux temps 1, 2 et 3 à la prédiction des résultats en décodage et en compréhension de la lecture. Pour ce faire, nous avons préalablement réalisé des ACP à l'aide des données des différentes épreuves. Les analyses ont permis de construire des phénomènes plus purs et ont mis en relief des tendances claires et cohérentes à travers les tranches de temps. Nous avons regroupé les épreuves et construit deux variables distinctes et orthogonales, représentant la MCT et le vocabulaire, que nous avons soumises à des analyses de régression, avec, comme variables prédites, la compréhension en lecture et le décodage.

5.2.1 Le vocabulaire en lien avec la compréhension en lecture

Nous avons vu que seuls les modèles des temps 1 et 3 impliquant la compréhension en lecture comme variable dépendante en temps final ont généré des modèles significatifs. Les résultats des analyses de régression, effectuées à la suite des analyses en composantes principales, indiquent que le vocabulaire a contribué significativement à la compréhension en lecture, plus que l'élément mémoire.

Les modèles des temps 1 et 3 mettent en relation les résultats des élèves de 5 et 7 ans avec leurs performances en décodage et en compréhension en lecture à 9 ou 10 ans. Le modèle du temps 1, formé des axes vocabulaire et MCT, contribue à expliquer 19% de la variance dans les performances des élèves à une épreuve de compréhension à 9-10 ans. L'axe vocabulaire est la seule variable significative du modèle. Ce résultat est compatible avec celui de Anderson et Freebody (1981), qui avaient conclu qu'un vocabulaire étendu tôt dans le développement de l'enfant était gage de bonnes performances ultérieures en lecture. Notre modèle de régression du temps 3 impliquant les épreuves comme variables prédictives explique 48,8% de la variance en compréhension de la lecture en temps final. Ces résultats concordent avec la longue liste d'articles ayant conclu au lien notable entre habiletés en vocabulaire et compréhension en lecture (Lervåg et Aukrust, 2010 ; Qian, 2002 ; Ricketts *et al.*, 2007 ; Shiotsu et Weir, 2007 ; Tunmer et Chapman, 2012).

Ainsi, nous avons observé que les R^2 des modèles impliquant les variables latentes MCT et vocabulaire ont été renforcés dans le temps : le modèle du temps 3 explique 30 % de plus de variance que le modèle du temps 1 – les R^2 sont passés de 0,19 à 0,49. Cette observation du renforcement des R^2 peut aussi servir une position inverse, ou complémentaire. Il est ainsi possible et même probable que les bons compreneurs

rencontrent plus de textes et que leur automatisation progressive du décodage leur permette d'accorder plus de ressources au vocabulaire. Ce vocabulaire, acquis grâce à la lecture, leur permet alors d'améliorer leurs lectures subséquentes. Cette position laisse sous-entendre que les deux construits s'influencent l'un l'autre (Nagy, 2005). Nos résultats ne nous permettent donc pas de trancher en faveur d'une position plutôt que l'autre.

Nous avons mentionné que la corrélation qui lie le vocabulaire à la compréhension est indéniable et attestée, mais demeure mal expliquée. Ceci est attribuable à la complexité de ce construit théorique et du flou qui entoure les processus qui le construisent. Une meilleure modélisation de ce que constitue la connaissance d'un mot est souhaitable. Une démarche expérimentale bonifiée sur le vocabulaire et sa relation avec la compréhension permettrait de vérifier si une même habileté est à la source des constantes corrélations qui sont observées. Certains auteurs ont suggéré, par exemple, que le QI et les habiletés en vocabulaire se nourrissent l'une l'autre. D'ailleurs, les tests d'intelligence incluent une mesure verbale ou de connaissance de mots ou de concepts (Cain *et al.*, 2004b). La vision capacitaire des processus cognitifs amenée par Daneman et Carpenter (1980) laisse envisager une piste afin de déterminer ce qui permet de construire des inférences et de traiter de nouvelles informations à partir de matériau linguistique, qu'il soit écrit ou oral.

5.2.2 Le vocabulaire en lien avec le décodage

Penchons-nous maintenant sur le lien qu'entretiennent le vocabulaire et le décodage. Dans le présent mémoire, le lien entre les habiletés en vocabulaire et le décodage n'a pas pu être établi. Perfetti (1985) avait proposé que les habiletés en vocabulaire et les performances en décodage étaient corrélées, tant chez les enfants que chez les adultes. Il stipulait, par ailleurs, que la compréhension en lecture, en tant que variable

prédite, était en partie fonction des habiletés en décodage et en vocabulaire. Ce lien a été prouvé dans de récentes études, notamment celle de Tunmer et Chapman (2012). Beck et McKeown (1991) (cité dans Aarnoutse *et al.*, 2001) avaient, quant à eux, suggéré que tant le décodage que la compréhension étaient fonction de l'étendue du vocabulaire.

A priori, il n'y a pas de raison pour qu'il y ait un lien entre le décodage et le vocabulaire. Nous avons préalablement défini le décodage comme une activité de bas niveau n'impliquant qu'un processus de conversion graphème-phonème. De plus, nous avons inscrit ce processus à l'intérieur du modèle DRC (Coltheart *et al.*, 2001), qui prévoit que l'accès au sens à partir du matériau écrit se fasse soit par une route lexicale, impliquant le vocabulaire, soit par une route impliquant le recodage des graphèmes en phonèmes. Notre conception du décodage anticipait les résultats nuls que nous observons dans notre étude entre celui-ci et le vocabulaire.

5.2.3 La MCT en lien avec la compréhension en lecture

La variable MCT n'a pas contribué de façon significative à la performance en compréhension au temps 1, même si le modèle était significatif. Ce résultat n'est pas inédit. Daneman et Carpenter (1980) rapportaient que les mesures de MCT n'étaient que faiblement ou même nullement corrélées dans plusieurs études. Ces résultats sont par ailleurs comparables à d'autres travaux (Cain *et al.*, 2004a ; Daneman et Carpenter, 1980 ; Swanson et Berninger, 1995). Le modèle du temps 2, quant à lui, ne s'est pas avéré significatif.

Comme nous l'avons vu, le modèle de régression du temps 3 a affiché une contribution négative de la variable latente MCT. Ce résultat infirme l'une de nos hypothèses de recherche, qui stipulait que la MCT, en tant que construit permettant le

stockage et la manipulation d'informations, exercerait une influence positive sur les performances en compréhension d'un point de vue longitudinal. Nos résultats ne nous permettent pas d'affirmer que les élèves qui réussissent moins bien à stocker à court terme des informations verbales sont ultimement de moins bons lecteurs. La MCT, telle que construite par les ACP, contribue significativement à expliquer une part de variance en compréhension ($\beta = -0,293$). Ainsi, un grand nombre d'élèves forts en MCT au temps 3 affichent des résultats moins bons en compréhension au temps final. Il nous faut nous demander pourquoi l'influence de la MCT est négative au temps 3.

Revenons à nos quatre épreuves de MCT des temps 1, 2 et 3, soit la répétition de phrases, la répétition de chiffres, la répétition de mots rares et la répétition de rythmes. Toutes comportent des processus de stockage et de récupération des informations en mémoire. Cependant, les épreuves diffèrent beaucoup sur le plan des éléments qui sont manipulés. L'épreuve de répétition de rythmes est particulièrement dissemblable ici, en ce sens que le matériau n'est pas linguistique. Cette épreuve n'a eu lieu qu'au temps 3. De plus, un coup d'œil au graphique des ACP (figure 4.3) nous permet de réaliser que l'épreuve construit massivement le facteur MCT au temps 3. La variable latente qui est construite par les épreuves au temps 3 est donc singulièrement différente de celles qui sont construites aux temps 1 et 2. L'inclusion de l'épreuve de répétition de rythmes dans les analyses de régression est peut-être en partie responsable de nos résultats surprenants, où la variable MCT contribue négativement à prédire les résultats en compréhension en lecture. Il se peut aussi que l'absence de l'épreuve de répétition de phrases dans la construction de la variable MCT du temps 3 ait fait la différence. La perte d'une tâche de répétition d'éléments de nature linguistique au profit d'une tâche de répétition d'éléments de nature non-linguistique est probablement la cause de ce que nous observons au temps 3.

L'influence négligeable que la MCT exerce sur la compréhension nous incite à nous aligner sur la position des dernières recherches sur l'acquisition de la lecture et les habiletés cognitives nécessaires à la lecture et à considérer que la MdT est bel et bien l'entité que l'on devrait chercher à mesurer si l'on veut prédire les performances en compréhension. Le caractère dynamique des processus qui sont en jeu dans la mémoire de travail explique que ce système soit plus à même d'expliquer la variation dans les performances en compréhension. Oakhill (1988, p. 22) résume bien cet état de choses : « *Understanding is not just the passive reception of information from a text, but the active putting together of a message, using several types of information* ».

Les différentes mesures de la MCT et l'ambiguïté entretenue par nombre d'études se consacrant à la MdT ou à la MCT contribuent à brouiller les pistes quant aux liens à tisser entre MCT et compréhension en lecture. Il est possible que les études ayant uni MCT et compréhension aient en fait mesuré certaines ou plusieurs des parties constitutives de la MCT verbale, comme la boucle phonologique (Kibby, 2009). La MCT constitue ainsi plutôt un sous-ensemble d'un processus plus général de traitement phonologique, qui influence la lecture. Les résultats présentés accréditent l'adoption d'une conception plus active d'un système de mémoire fonctionnant avec des sous-systèmes spécifiques pour rendre compte de la variance en compréhension (Cain *et al.*, 2004a ; Van der Linden, Schelstraete, Hupet et Desmette, 1995). La version statique et unitaire de ce système, instrumentalisée par des tâches de simple répétition et, qui plus est, de nature diverses et hétéroclites, ne parvient pas à expliquer les performances en compréhension.

5.2.4 La MCT en lien avec le décodage

Les analyses de régression impliquant la MCT et le décodage ne se sont pas avérées significatives. Cet état de choses peut s'expliquer par les natures différentes des processus mis en œuvre dans chacune des habiletés : le dynamisme inhérent à l'objet de mémoire capacitaire est peu compatible avec l'opération automatique et basique qu'est le décodage. Les épreuves de décodage, soit la lecture de mots et de non-mots en une minute, n'impliquaient pas de saturation de la MCT. En effet, le nombre maximal de syllabes présentes dans chacun des mots est de quatre. Ce chiffre n'atteint même pas la limite inférieure du « chiffre magique » de 7 plus ou moins 2 éléments de (Miller, 1956). De plus, les épreuves de décodage ne comportaient pas de double-tâche, comme on peut en trouver dans les épreuves de mémoire de travail. Ainsi, aucune surcharge cognitive n'a été observée.

5.3 Considérations d'ordre méthodologique

5.3.1 Considérations méthodologiques particulières

Les différentes analyses statistiques effectuées ont mis au jour un problème en temps 2. En effet, nous avons remarqué certaines corrélations étonnantes, notamment dans les corrélations entre différentes variables construites – par exemple, celle unissant la variable MCT au T1 et celle du T2 ($r = -0,477$). Les analyses de régression en temps 2 étaient régulièrement moins significatives et moins parlantes. En revenant sur certaines données mesurées, nous avons remarqué que les notes en T2 étaient plus basses qu'en T1 et T3. Pourtant, le patron usuel, et celui que nous nous attendions à observer, en est un où on voit une progression globale de l'ensemble des élèves au fur et à mesure de leur avancement en âge et en niveau scolaire. Nous avons voulu savoir si le patron auquel nous faisons face était un effet du progrès général ou individuel. Nous avons procédé à de nouvelles analyses sur les mesures qui semblaient causer problème, soit les épreuves de mémoire. Nous avons remarqué que les enfants performaient moins bien en T2, particulièrement à l'épreuve de répétition de mots

rare. En analysant les données brutes, nous avons repéré les items problématiques, qui correspondaient à la tâche appelée « série phonologique » dans le NÉEL. Ces items avaient la particularité de contenir des *clusters* phonologiques, c'est-à-dire qu'ils contenaient des syllabes avec des structures de type CCV (*perspicacité*) ou CCCV (*obstruction*). La répétition de mots contenant ce type de constructions peut représenter une certaine difficulté. Nous nous expliquons toutefois très mal pourquoi un tel patron n'est visible qu'au T2. Pourquoi les enfants seraient-ils moins performants à cette même tâche quelques mois après leur première passation? Il peut certainement s'agir également d'un effet de l'expérimentateur, dans la mesure où cette tâche est la première à laquelle les enfants sont exposés, et ce, lors des passations en temps 1, 2 et 3. Nous nous sommes résignée à ne pas pouvoir expliquer l'effet observé. Nous avons donc fait le choix de garder les données intactes, dans le but de les analyser dans leur ensemble et telles quelles sont utilisées dans le NÉEL.

5.3.2 Considérations méthodologiques générales

Il faut ici convenir du fait que nous avons discuté, dans le cadre de ce mémoire, des concepts de MCT et de vocabulaire, à partir de tâches ne représentant qu'imparfaitement ces deux construits théoriques. Les épreuves du NÉEL ont été conçues pour diagnostiquer des problèmes dans le langage oral des enfants. Elles n'ont pas été conçues pour circonscrire fidèlement des habiletés spécifiques. Nous constatons qu'il existe une contradiction entre les visées diagnostiques et explicatives d'une recherche. D'un côté, on cherche à élargir le champ des données, en tentant de trouver des épreuves qui mobilisent des habiletés de natures diverses et donc, mesurent beaucoup, d'un seul coup. L'épreuve de répétition de phrases constitue un bon exemple de ce type de tests pluriels et efficaces. Rappelons que cette épreuve a représenté un bon dépisteur des problèmes ultérieurs en compréhension de la lecture dans notre étude. Cependant, cette épreuve, qui met en jeu une variété d'habiletés, ne

permet pas d'affiner notre compréhension des phénomènes en jeu dans la réussite scolaire ou les performances en lecture. Ainsi, le fait de mesurer une tâche de haut niveau, de nature complexe, à l'aide d'une autre tâche de haut niveau, presque aussi complexe, ne permet pas de comprendre les habiletés exigées par la lecture. Les travaux aux ambitions explicatives ne sont pas bien servis par ce genre d'épreuves; ils le sont plus par les outils de mesure plus « purs », qui tentent de circonscrire une seule habileté. Malheureusement, dans le présent mémoire, les tâches qui définissent la MCT, comme les mesures classiques de répétition de mots rares et de répétition de chiffres, n'ont pas contribué significativement aux mesures de compréhension en lecture et de décodage et ne nous ont donc pas permis d'attribuer de l'importance à ce que représente la MCT pour les performances en lecture dans la population étudiée.

Dans un autre ordre d'idées, malgré que nos résultats soient cohérents avec ceux d'autres études concernant le pouvoir prédictif du vocabulaire pour la compréhension en lecture, ceux-ci doivent être interprétés avec prudence. Le fait que les performances en compréhension en lecture aient été mesurées à l'aide d'une seule épreuve, composée de 9 textes et, pour chacun d'eux, d'une question à choix de réponse, est peut-être problématique. Il se peut que la tâche ait échoué à représenter la réelle capacité de compréhension des enfants évalués. Le codage de l'épreuve, qui était effectué dans une logique ternaire (0, 1 ou 2 points par question), n'a peut-être pas permis de mesurer une entité complexe comme les performances en compréhension de la lecture. De plus, l'épreuve a été construite massivement pour représenter la notion d'inférence. Dans la mesure où la compréhension se fait grâce à un ensemble d'habiletés, nous constatons qu'un plus grand groupe d'épreuves auraient été nécessaires afin d'en arriver à une mesure composite et représentative de l'activité de compréhension.

CONCLUSION

Nous avons voulu explorer les contributions relatives de variables mesurées, tirées d'épreuves d'une batterie d'évaluation du langage oral, et de variables latentes, elles-mêmes construites à partir de ces mesures, dans la prédiction des performances en lecture. Nous nous sommes intéressée plus particulièrement à une communauté autochtone vivant d'importants problèmes de décrochage scolaire et accusant un retard sur le plan de la lecture, par rapport aux jeunes non-autochtones. Nous avons, dans un premier temps, identifié les épreuves nous permettant de mieux prédire les performances des enfants de cette communauté en lecture. Dans un deuxième temps, nous nous sommes penchée sur l'importance relative des habiletés ressortant des épreuves, soit la MCT et le vocabulaire, dans les performances en lecture.

D'un point de vue diagnostique, les résultats des analyses suggèrent que ce sont les épreuves de répétition de phrases et de vocabulaire (expression et réception) qui sont les plus à même de prédire efficacement les performances ultérieures en compréhension de la lecture. Ces épreuves font toutes fonctionner, à différents degrés, des habiletés en vocabulaire. Les épreuves impliquant la répétition d'éléments ne permettent pas de prédire les performances finales en lecture : les liens entre MCT et lecture sont nuls. Le caractère statique des épreuves représentant la MCT est probablement la cause de ce constat. Nous nous permettons de suggérer qu'il serait probablement bénéfique d'identifier les enfants à risque à l'aide d'épreuves évaluant plutôt la MdT, à l'aide d'épreuves adaptées (Daneman et Carpenter, 1980 ; Delafoy et Ehrlich, 1990 ; Van der Linden *et al.*, 1995).

Nos résultats concernant les variables latentes concordent avec ceux ayant trait à nos variables mesurées. D'un point de vue explicatif, nos analyses nous permettent de

conclure que les habiletés en vocabulaire contribuent à expliquer les performances en compréhension de la lecture. Le lien qu'entretiennent ces deux construits théoriques est extrêmement complexe (Baumann, 2005). L'analyse des processus cognitifs à la source du lien vocabulaire-compréhension est une entreprise laborieuse. Notre étude s'inscrit pourtant parmi celles qui affichent une position « instrumentaliste », où les individus connaissant beaucoup de mots sont de bons « compreneurs » (Beck, McKeown et Omanson, 1987 ; cité dans Baumann, 2005, p.127). Nos résultats n'excluent toutefois pas l'hypothèse que les deux construits soient tous deux le reflet d'une habileté plus générale, ou encore, qu'ils se construisent l'un l'autre, comme le propose Nagy (2005).

Considérons maintenant la contribution relative de la variable latente MCT à la performance en compréhension de la lecture. Rappelons que les différentes épreuves de mémoire consistaient en un stockage passif de l'information. Daneman et Carpenter, auteurs de l'épreuve d'empan de lecture (MdT), ont parié que le lien qui unit la compréhension en lecture et la MdT est tributaire de la nature des deux habiletés. En effet, les deux types de tâches mobilisent des processus langagiers (*domain-specific*). Cette relation s'explique probablement par le fait que la compréhension, processus de haut niveau s'il en est, requiert non seulement le maintien d'informations, mais aussi le traitement de ces informations, par exemple en les mettant en lien avec des expériences passées ou toute autre opération relevant de la construction d'un modèle mental, comme les inférences. En ce sens, la MdT, qui a un caractère dynamique, correspond mieux aux processus sous-jacents de la compréhension. Il n'est pas étonnant que la MCT ne soit pas parvenue à expliquer avec succès la performance à une habileté langagière telle que la compréhension en lecture. La MdT est probablement plus à même d'expliquer la variation dans les performances en lecture. D'ailleurs, depuis les années 1980, les travaux portant sur les processus cognitifs de haut niveau, comme la lecture et l'écriture, par exemple, ont délaissé la MCT au profit de la MdT (Baddeley, 1986 ; Daneman et Carpenter,

1980 ; Just et Carpenter, 1992). Comme l'explique McCutchen (1996), les processus dynamiques en jeu dans la MdT correspondent mieux à ces activités.

Il aurait été intéressant de procéder à une actualisation du MSL, tel que conçu par Gough et Tunmer (1986) et Hoover et Gough (1990), à l'aide des données dont nous disposons. Ceci constitue une piste de recherche à envisager. Par exemple, de vérifier l'importance du décodage dans la population à l'étude, non pas en tant que composante de la lecture, telle que nous l'avons fait dans la présente recherche, mais bien comme prédicteur des performances ultérieures en compréhension de la lecture, à l'instar de plusieurs (Joshi et Aaron, 2000 ; Netten *et al.*, 2011 ; Tunmer et Chapman, 2012). Une compréhension accrue des habiletés mises en jeu dans la lecture, notamment le décodage, est nécessaire afin d'arriver à un diagnostic et une prise en charge efficaces, notamment dans des communautés aux prises avec des conditions socioéconomiques particulières.

Si nos résultats confirment le lien notable entre habiletés en vocabulaire et compréhension en lecture, et ce, longitudinalement, ils montrent aussi qu'une mesure statique de la mémoire (la MCT) n'est pas un bon prédicteur des performances en compréhension. En outre, ce mémoire illustre la contradiction entre les visées diagnostiques et explicatives d'une recherche.

ANNEXE A

ÉPREUVES DU TEMPS FINAL

- A.1 Feuille de l'expérimentatrice – Épreuve de lecture de mots en une minute (temps final)
- A.2 Feuille de l'expérimentatrice – Épreuve de lecture de non-mots en une minute (temps final)
- A.3 Épreuve de compréhension (temps final)

ANNEXE A.1

ÉPREUVE DE LECTURE DE MOTS EN UNE MINUTE

L'épreuve a été adaptée de (Khomsi, 1999).

Consignes à dire à l'enfant :

Sur la feuille, il y a une liste de mots. Je veux que tu les lises, le mieux et le plus vite que tu peux, à haute voix. Tu ne pourras pas tous les lire, parce que je vais t'arrêter à un moment pour pouvoir changer d'exercice. Liste de mots :

| | | | |
|---------|-----------|-------------|--------------|
| il | poisson | devient | travail |
| un | herbe | mardi | yacht |
| le | ped | drôle | animal |
| lui | chemin | histoire | surprise |
| nu | joie | glacé | doigt |
| os | phare | derrière | sommeil |
| et | ciel | stylo | réunir |
| fil | pharmacie | magasin | briller |
| bol | enveloppe | cueillir | guêpe |
| tous | éléphant | compagnon | feuillage |
| mur | trésor | chronomètre | gymnastique |
| sac | choc | continuent | fauteuil |
| clé | fier | soleil | obtenir |
| est | nouveau | pharmacie | jusque |
| rue | expert | enveloppe | expliquer |
| petit | sportif | éléphant | magnifique |
| camion | coup | trésor | spécial |
| nom | chorale | choc | fenouil |
| acide | bruit | fier | gourmand |
| film | monsieur | nouveau | chirurgien |
| grand | agir | expert | prévenir |
| mars | cousin | sportif | assiette |
| jaune | unir | merveilleux | dangereux |
| parc | naïf | odeur | observer |
| sept | curieux | plongeant | obscurité |
| bloc | venir | parent | descendre |
| faim | départ | idéal | installeront |
| compter | écho | obéir | inquiétude |
| | | griller | |

ANNEXE A.2

ÉPREUVE DE LECTURE DE NON-MOTS EN UNE MINUTE

L'épreuve a été adaptée de BALE (2010).

Consignes à dire à l'enfant :

Sur la feuille, il y a une liste de mots. Ces mots ne sont pas des vrais mots en français. Ça ne fait rien, je veux que tu les lises quand même, le mieux et le plus vite que tu peux, à haute voix. Tu ne pourras pas tous les lire, parce que je vais t'arrêter à un moment pour pouvoir changer d'exercice.

Liste de mots :

| | | |
|-----------|-----------|-----------|
| sande | vatrice | cin |
| chon | pisal | taleau |
| givor | bertale | gantin |
| bondeuse | pacirande | racin |
| sule | anchovée | fachou |
| toir | agante | gontra |
| mic | courlone | copage |
| taubage | stipe | bartin |
| fudin | torac | datoir |
| esan | cassine | majon |
| trane | bate | nagule |
| tagin | coginte | savette |
| splindron | abranise | bracho |
| modan | glon | famir |
| tandir | ceau | poulan |
| taparelle | gal | flocachin |
| abindeur | deire | abranise |
| gental | acu | verdulin |
| ontage | gean | abritel |
| rac | isan | scropalte |
| gavin | chan | gilone |
| caldon | mythe | gordivet |
| rigende | banton | siropage |
| plour | thénou | corabone |
| | | pontaneur |

ANNEXE A.3

ÉPREUVE DE COMPRÉHENSION EN LECTURE

L'épreuve de compréhension en lecture a été construite à partir des textes fournis dans l'ouvrage de Foucambert (2003) et des articles de Ouellet (2012), de Ouellet (2013) et de Morissette (2012), ces derniers provenant de la version en ligne du magazine *Les Débrouillards*.

Consignes à dire à la classe :

Je vais distribuer à tout le monde des petits cahiers dans quelques instants. Sur chaque page, côté recto, il y a un texte. Tu peux le lire autant de fois que tu veux. Tu peux aussi mettre des étoiles au bas de la page pour indiquer le nombre de fois où tu as lu le texte. Une fois que tu l'as lu suffisamment et que tu as l'impression de l'avoir compris, tu peux tourner la page. Au verso de la page, il y a une question qui porte sur le texte que tu viens de lire. Tu ne peux pas retourner au texte une fois la page tournée. Tu dois choisir la meilleure réponse, parmi les trois réponses possibles, en te fiant à ce que tu as retenu du texte. Tu fais la même chose pour le reste des textes. Tu ramènes le cahier à l'avant de la classe après avoir répondu à chaque question.

TEXTE 1

Un bébé de vingt mois jouait sur la plage. Vers 6 heures du soir, le bébé a été attaqué par un berger allemand de grande taille. Le chien a mordu l'enfant à l'arrière de la tête. En tentant de protéger son fils, la mère a aussi été mordue au bras gauche. Pourtant, la présence des chiens sur les plages est interdite dans cette région.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) d'un homme qui promenait son chien
- B) d'un enfant qui jouait sur une plage
- C) d'un bébé qui a été mordu par un chien

TEXTE 2

Le lézard est dans une grande cage. La scène est éclairée d'une pâle lumière rouge. Un serpent est introduit doucement. Les hommes veulent savoir comment le lézard se défend. Dès qu'il aperçoit le serpent, le lézard soulève sa queue et l'agite. Il va même la secouer sous le nez du serpent. Trois fois sur dix, le serpent attrape la queue du lézard. Elle se détache tout de suite et cela permet à son propriétaire de s'enfuir pendant que le serpent la mange.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) du moyen de défense du lézard
- B) de l'agressivité du serpent
- C) des habitudes alimentaires du lézard

TEXTE 3

Une fillette de quatre ans a été retrouvée saine et sauve après avoir sauté du quatrième étage d'un immeuble pour échapper à un incendie. La petite Christelle se trouvait seule en compagnie de son petit frère. Elle a eu peur quand le feu s'est déclaré dans une chambre de l'appartement. Elle a alors sauté par la fenêtre. Heureusement, la petite s'est retrouvée sur la pelouse sans une égratignure, quinze mètres plus bas. L'incendie a rapidement été éteint. Le petit frère de la fillette a été retrouvé lui aussi sain et sauf.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) de la manière dont une fillette a échappé à un incendie
- B) d'une fillette qui a sauté par une fenêtre de son immeuble
- C) d'une fillette qui a été sauvée par des pompiers

TEXTE 4

Jack Andraka, 15 ans, s'est intéressé au cancer du pancréas lorsque le frère d'un ami a été emporté par cette maladie. En consultant des articles scientifiques sur Internet, ce jeune Américain a eu l'idée de développer un nouveau test pour déceler la présence de la maladie.

En plus d'être fiable à 90 %, sa technique est huit fois plus rapide et beaucoup moins coûteuse que les tests actuels. Les efforts de Jack ont été récompensés, puisqu'il a

remporté 100 000 \$ lors d'un concours scientifique. Le jeune chercheur espère rendre son invention accessible à tous et l'adapter pour diagnostiquer d'autres maladies. Belle carrière en vue, Jack!

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) d'un jeune Américain qui a gagné un concours scientifique
- B) d'un jeune homme décédé d'un accident de voiture
- C) d'un jeune qui a mis au point un test de dépistage du cancer

TEXTE 5

Pierre et André sont partis de Cap Breton à bord d'un hélicoptère et ont atteint l'île d'Anticosti. Ils ont battu, en dix heures de vol, la meilleure performance de traversée maritime. L'hélicoptère des deux pilotes, qui devait se poser sur l'héliport de l'île, est tombé en mer tout près du point d'arrivée. Ils avaient tenté de se poser trois fois, sans succès, sur une petite piste de 14 mètres de large et 125 mètres de long. À cause d'un manque de place, l'équipage a dû à chaque fois remettre les gaz pour repartir. Les pilotes ont finalement préféré terminer leur voyage sur l'eau.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) d'une traversée du golfe du Saint-Laurent en bateau à moteur
- B) d'une arrivée mouvementée
- C) d'une traversée maritime record

TEXTE 6

Pour la première fois, un ours brun a été vu en train de se servir d'une pierre pour nettoyer son pelage. Il a été observé pendant une longue minute alors qu'il prenait son bain dans un ruisseau, en Alaska. Un chercheur anglais a eu la chance de voir ce comportement jusqu'alors inconnu chez l'ours.

On savait déjà que l'ours, en proie à des démangeaisons, se gratte le dos contre un arbre ou utilise ses griffes. Mais un nettoyage à la pierre, c'est une découverte! Pourquoi est-ce si surprenant? Car généralement, ce sont les primates (dont le singe et l'homme) qui utilisent des outils pour effectuer leurs tâches quotidiennes. L'ours devient ainsi l'un des rares animaux non primates connus à utiliser un outil pour se laver.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) d'un chercheur anglais en Alaska
- B) du mauvais caractère de l'ours brun
- C) de la technique de nettoyage de l'ours brun

TEXTE 7

L'équipage du brise-glace canadien *Shediac* a aidé un millier de baleines à regagner la haute mer à la sortie du golfe Saint-Laurent. Comment ? En diffusant de la musique classique ! Ces baleines, longues de 3 à 4 mètres, ont failli être victimes de leur gourmandise. Elles s'étaient aventurées sur les hauts fonds entre Baie-Comeau et Matane en poursuivant un immense banc de poissons. Par malheur, elles s'étaient retrouvées prisonnières des glaces.

Les autorités canadiennes, ne souhaitant pas voir disparaître des baleines, ont décidé de tenter une opération de secours en creusant un canal de 20 kilomètres. Mais comment les attirer ? À bord, quelqu'un a eu l'idée de diffuser de la musique par haut-parleurs. Ainsi, les baleines ont pu être sauvées... Mais les chasseurs de baleines ne les épargneront pas.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) du sauvetage de baleines prisonnières des glaces
- B) de chasseurs de baleines
- C) du manque de poissons dans le fleuve du Saint-Laurent

TEXTE 9

Les éleveurs d'abeilles de l'Estrie risquent d'avoir une mauvaise surprise quand ils vont ouvrir leurs ruches. Non seulement ils n'y trouveront pas le miel espéré, mais ils y verront des abeilles affamées. Jean Charron, président du syndicat des apiculteurs de l'Estrie, tient à les prévenir. S'ils attendent quelques semaines avant de regarder les cadres, il sera peut-être trop tard : les abeilles seront mortes ou auront quitté la ruche. «Je n'en suis pas revenu moi-même quand j'ai ouvert mes ruches il y a quelques jours. Ça fait trente ans que je fais du miel et je n'ai jamais vu ça !», affirme Jean.

L'apiculteur explique ainsi la situation : «Le printemps a été froid, ensuite il a fait très

sec. Les fleurs n'ont pas produit autant de nectar que d'habitude. Les abeilles n'ont pas pu faire de provisions et il va falloir les nourrir avec du sirop (1 livre de sucre pour 1 litre d'eau). La récolte s'annonce catastrophique. Au Québec, la production sera inférieure à dix tonnes, contre une centaine habituellement».

Heureusement, les deux récoltes précédentes avaient été très bonnes.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) du mode de vie des abeilles
- B) des fleurs qui produisent peu de nectar
- C) de l'influence du climat sur la production de miel

TEXTE 10

En 2004, Claudia Garneau et Charlaïne Dalpé avaient 12 ans. Elles passaient leurs vacances à Grande-Vallée, en Gaspésie. Elles ont inséré un message dans une bouteille de plastique et l'ont lancée à la mer. Surprise! Un garçon de 9 ans vient de retrouver leur bouteille sur une plage d'Irlande. La bouteille a traversé l'Atlantique, parcourant plus de 4 000 km.

Quand Vik Millea a découvert le message, il a couru voir sa mère pour qu'elle l'aide à le déchiffrer. Comme il était rédigé en français, ils ont utilisé un outil de traduction sur Internet. Le garçon s'est empressé d'écrire à cette adresse courriel, mais elle n'était plus en fonction. Par chance, un ami de Charlaïne a lu un avis de recherche publié dans un journal montréalais et il a contacté la jeune fille. Le lendemain, une rencontre sur Internet s'est organisée entre Vik, Claudia et Charlaïne. Et pour terminer en beauté ce « conte de fée », une agence de tourisme a invité les deux jeunes femmes à séjourner en Irlande l'été prochain. Elles en profiteront pour rencontrer Vik et revoir la lettre qui a fait le tour des médias.

Choisis la meilleure réponse.

Pour toi, le texte parle surtout :

- A) de deux jeunes filles qui passent leurs vacances en Gaspésie
- B) d'une bouteille lancée à la mer qui a été retrouvée par un jeune Irlandais
- C) d'une lettre d'amour

ANNEXE B

SUPPLÉMENTS AUX ANALYSES

B.0 Comparaison des résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence

B.1 Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 1

B.2 Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 2

B.3 Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 3

ANNEXE B.0

Comparaison des résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence

Plusieurs documents gouvernementaux font état de l'écart sur le plan de la réussite scolaire entre les élèves autochtones et non-autochtones (voir notamment les rapports du CMEC (2008) et de la CDPDJ (2009)). Nous savons aussi que les enfants autochtones possèdent des faiblesses en lecture et en vocabulaire (Morris et Mackenzie, 2013). Nous avons voulu vérifier comment se traduisaient ces écarts dans les résultats des élèves de la communauté à l'étude par rapport à ceux d'enfants appartenant à une communauté avec ne présentant pas un profil socio-économique ou un contexte linguistique particuliers. Or les instruments de mesure de l'ensemble de la batterie NÉEL ont été utilisés sur une population d'élèves français d'écoles maternelles et élémentaires et les résultats de cette enquête sont disponibles dans la batterie d'évaluation. L'échantillon de référence a été construit afin de refléter la répartition des catégories professionnelles que l'on trouve dans la population française en général (Chevrie-Muller et Plaza, 2001), ce qui rend intéressants les parallèles.

Nous avons procédé à une comparaison systématique des résultats de ces deux groupes pour toutes les épreuves communes, en nous assurant que les échantillons étaient d'une même classe d'âge. Afin de vérifier si écart il y avait entre les résultats des deux groupes, nous avons fait une comparaison d'échantillons à l'aide de tests t de student bilatéral. Les représentations graphiques de ces analyses figurent dans les pages suivantes. Il est à noter que les regroupements effectués dans le cadre du mémoire et leurs libellés ne se retrouvent pas tel quels dans les graphiques qui suivent. En effet, nous avons dû séparer les épreuves de la même façon que le fait le manuel du NÉEL (Chevrie-Muller et Plaza, 2001).

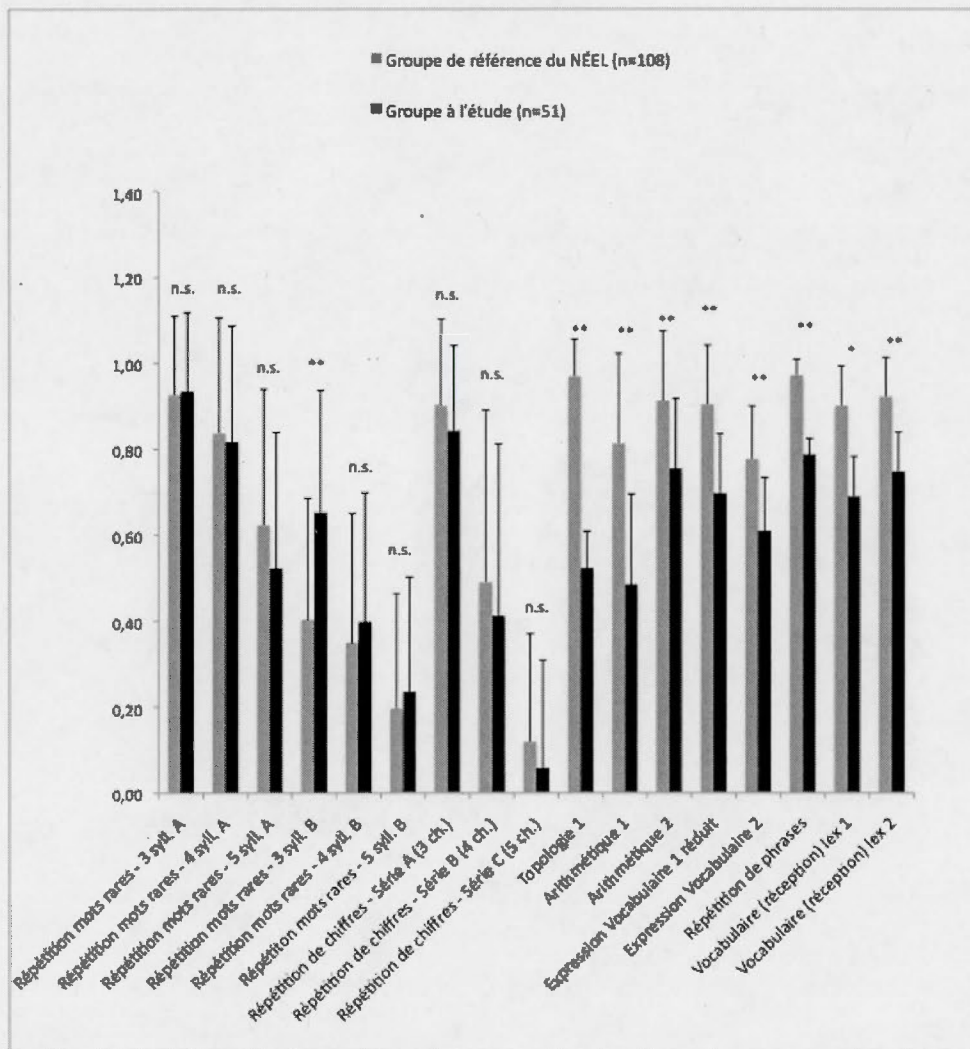
Notons que les analyses ont été réalisées malgré les effets plafond observés dans un grand nombre d'épreuves dans les résultats du groupe de référence du NÉEL ainsi que de la communauté à l'étude. Les résultats du groupe de référence doivent être analysés avec précaution, étant donné qu'ils n'ont pas été normalisés. Il n'en demeure pas moins que les graphiques présentés permettent de situer la communauté à l'étude par rapport à une communauté ne présentant pas les mêmes particularités.

En guise de discussion, remarquons que le groupe de la communauté à l'étude présente plusieurs résultats significativement plus faibles que le groupe de référence, surtout en ce qui a trait au vocabulaire. Ce patron est observable à tous les temps. On peut observer, dans certaines épreuves, des résultats à plus de deux écarts-types de la moyenne par rapport aux résultats du groupe de référence. Cette observation n'est cependant pas généralisable. Aux T1 et T2, elle ne touche que les épreuves topologie 1, de répétition de phrases et de vocabulaire (réception, lexique 1 et 2). Au T3, ce sont les épreuves de topologie 2 et 3 et vocabulaire (réception) lexique 1. On ne peut nier qu'il y a un écart entre les résultats des petits français par rapport aux jeunes de la communauté en ce qui a trait au vocabulaire. Cependant, les minuscules écarts-types que présentent les résultats des élèves de la communauté de référence, conséquence directe des effets plafond observés dans un grand nombre d'épreuves, exagèrent cet état de choses. Comme il n'y a pas vraiment de variation dans l'échantillon français, les écarts-types sont réduits. Ceux-ci ne nous informent pas vraiment, ni sur la communauté de référence, ni sur la communauté que nous étudions.

Globalement, les élèves la moyenne des résultats aux épreuves de la communauté à l'étude est inférieure à celle de la communauté de référence. Ce résultat global est probablement d'abord le reflet de la situation linguistique de la communauté, où, pour plusieurs enfants, le français constitue une L2, la langue apprise à l'école. L'écart témoigne probablement aussi des problèmes de pauvreté et d'exclusion qui vit la communauté. Nous constatons que les enfants font trop souvent les frais du manque d'exposition à la langue et d'un contexte social appauvri (Snow *et al.*, 1998).

ANNEXE B.1

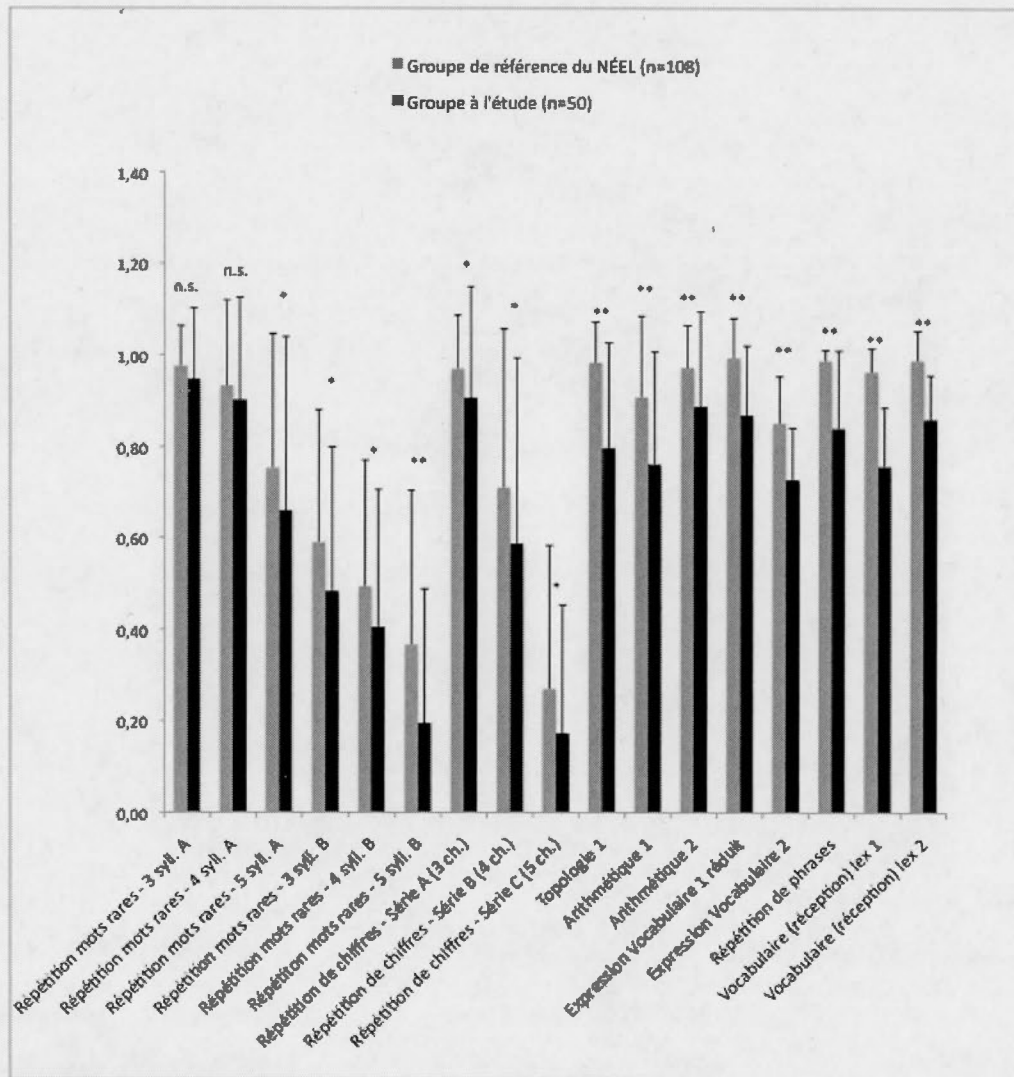
Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 1



Note : Forme P, 5 ans; * $p < 0,05$; ** $p < 0,01$

ANNEXE B.2

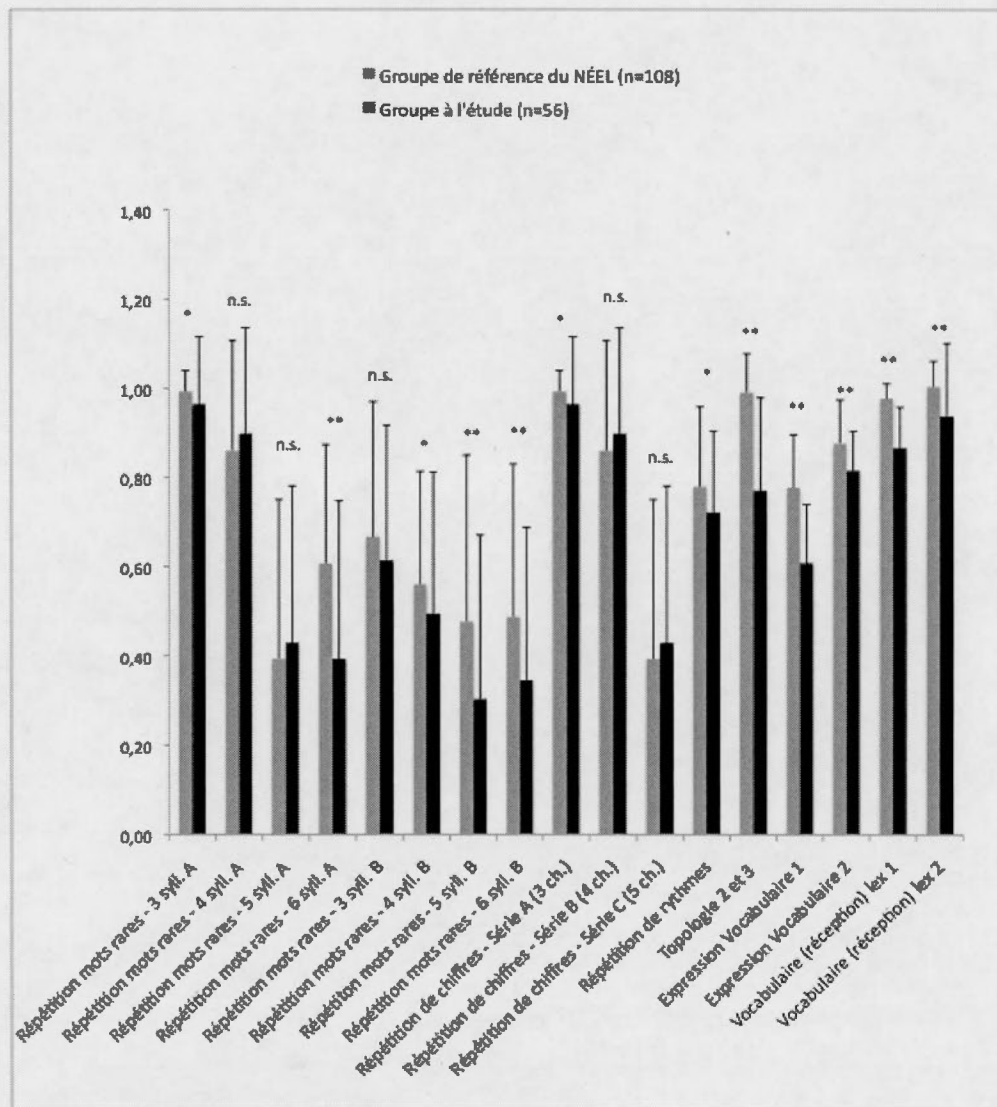
Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 2



Note : Forme P, 6 ans; * $p < 0,05$; ** $p < 0,01$

ANNEXE B.3

Résultats des épreuves du NÉEL pour la communauté à l'étude et le groupe de référence au temps 3



Note : Forme G, 7 ans; * p<0,05; ** p<0,01

BIBLIOGRAPHIE

- Aarnoutse, C. *et al.* (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. *Reading and Writing: An Interdisciplinary Journal*, 14, 61–89.
- Alamargot, D. et Chanquoy, L. (2002). Mémoire de travail et rédaction de textes : évolution des modèles et bilan des premiers travaux. *L'année psychologique*, 363-398.
- Alloway, T.P. et Gathercole, S.E. (2005). The role of sentence recall in reading and language skills of children with learning difficulties. *Learning and Individual Differences*, 15, 271–282.
- Anderson, R.C. et Freebody, P. (1981). Vocabulary knowledge. Dans Guthrie, J. T. (dir.), *Comprehension and teaching: Research reviews* (p. 77-117). Newark : DE: International Reading Association.
- April, N. et Bourret, A. (2004). *État de la situation sur le Syndrome d'Alcoolisation Foetale au Québec*. Québec: Institut national de santé publique du Québec.
- Archibald, L.M.D. et Gathercole, S.E. (2007). Nonword repetition and serial recall: Equivalent measures of verbal short-term memory? *Applied Psycholinguistics*, 28(04), 587–606.
- Atkinson, R. et Shiffrin, R. (1968). Human memory: A proposed system and its control processes. Dans Spence, K. et J. Spence (dir.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 2, p. 89-195). New York : Academic Press.
- Atkinson, R. et Shiffrin, R. (1971). *The control processes of short-term memory*. Stanford University : Institute for Mathematical Studies in the Social Sciences.
- August, D. et Shanahan, T. (2006). *Developing literacy in second-language learners : report of the National Literacy Panel on Language Minority Children and Youth*. : Mahwah, N.J. : Lawrence Erlbaum.

- Baddeley, A. (1986). *Working memory*. Oxford : Oxford University Press.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36(3), 189-208.
- Baddeley, A. et Hitch, G. (1974). Working Memory. Dans Gordon, H. B. (dir.), *Psychology of Learning and Motivation* (Vol. Volume 8, p. 47-89) : Academic Press.
- Badian, N.A. (1990). Background Factors and Preschool Test Scores as Predictors of Reading: a Nine-year Longitudinal Study. *Reading and Writing: An Interdisciplinary Journal*, 2, 307-326.
- BALE. (2010). *Batterie Analytique du Langage Écrit, partie 2 : Outils*. UPMF - Grenoble.
- Baumann, J.F. Maloch, B., J. V. Hoffman, D. L. Schallert, C. M. Fairbanks et J. Worthy (dir.). (2005). *Vocabulary-Comprehension Relationships. Fifty-fourth yearbook of the National Reading Conference*, Actes du colloque, 2005, Oak Creek, WI: National Reading Conference
- Beck, I.L. et McKeown, M.G. (1991). Conditions of vocabulary acquisition. Dans Barr, R., M. L. Kamil, P. B. Mosenthal et P. D. Pearson (dir.), *Handbook of reading research, Vol. 2* (p. 789-814). Hillsdale, NJ, England : Lawrence Erlbaum Associates, Inc.
- Bertrand, D. *et al.* (2010). Efficacité, sensibilité, spécificité : comparaison de différents tests de lecture. *L'année psychologique*, 110, 299-320.
- Bodson, H. (2013). *L'évaluation de la progression de la discrimination des phonèmes du français et de la conscience phonologique chez les enfants innus de la maternelle*. (Mémoire de maîtrise). Université du Québec à Montréal, Montréal.
- Bogaards, P. (2000). Testing L2 vocabulary knowledge at a high level: the case of the Euralex French Tests. *Applied Linguistics*, 21(4), 490-516.
- Butler, S.R. *et al.* (1985). Seven-year longitudinal study of the early prediction of reading achievement. *Journal of Educational Psychology*, 77(3), 349-361.

- Cain, K. et Oakhill, J. (2011). Matthew Effects in young readers: reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, 44(5), 431-443.
- Cain, K., Oakhill, J. et Bryant, P. (2004a). Children's Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills. *Journal of Educational Psychology*, 96(1), 31-42.
- Cain, K., Oakhill, J. et Lemmon, K. (2004b). Individual Differences in the Inference of Word Meanings From Context: The Influence of Reading Comprehension, Vocabulary Knowledge, and Memory Capacity. *Journal of Educational Psychology*, 96(4), 671-681.
- Campion, N. et Rossi, J.-P. (1999). Inférences et compréhension de texte. *L'année psychologique*, 99(3), 493-527.
- Carretti, B. et al. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, 19(2), 246-251.
- Carricano, M., Poujol, F. et Bertrandias, L. (2010). *Analyse de données avec SPSS®*, 2e édition. Paris, France : Pearson Education.
- Carroll, J.B. (1993). *Human cognitive abilities a survey of factor-analytic studies*. Cambridge, Angleterre : Cambridge University Press.
- Catts, H.W., Fey, M.E. et Zhang, X. (2002). A Longitudinal Investigation of Reading Outcomes in Children With Language Impairments. *Journal of Speech, Language, and Hearing Research*, 45, 1142-1157.
- Catts, H.W. et al. (2001). Estimating the Risk of Future Reading difficulties in Kindergarten children : a Research-Based Model and Its Clinical Implementation. *Language, Speech, and Hearing Services in Schools*, 32, 38-50.
- CDPDJ. (2009). *Mythes et réalités sur les peuples autochtones*, 2e édition. Québec: Commission des droits de la personne et des droits de la jeunesse.
- Chevrie-Muller, C. et Plaza, M. (2001). *Nouvelles épreuves pour l'examen du langage : bilan complet des capacités langagières et cognitives de l'enfant*. Paris : Les Éditions du Centre de Psychologie Appliquée.
- CMEC. (2008). *L'Éducation au Canada – Horizon 2020*. Toronto: Conseil des ministres de l'Éducation (Canada).

- Cobb, T. (2013). *The Compleat Lexical Tutor, v.4* Montréal. Récupéré de lextutor.ca
- Cohen, J. et al. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, 3rd Edition*. Mahwah, New Jersey : Lawrence Erlbaum Associates.
- Coltheart, M. (2004). Are there lexicons? [Review]. *The Quarterly Journal of Experimental Psychology, 57*(7), 1153-1171.
- Coltheart, M. et al. (2001). DRC: A Dual Rote Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review, 108*(1), 204-256.
- Conway, A.R.A. et al. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786.
- Daneman, M. et Carpenter, P.A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.
- De Beni, R. et Palladino, P. (2000). Intrusion errors in working memory tasks : Are they related to reading comprehension ability? *Learning and Individual Differences, 12*, 131-143.
- Delafoy, M. et Ehrlich, M.-F. (1990). La mémoire de travail : structure, fonctionnement, capacité. *L'année psychologique, 90*(3), 403-427.
- Desrochers, A. et Berger, M.J. (2011). *L'évaluation de la littératie*. Ottawa : Presses de l'Université d'Ottawa.
- Desrochers, A., Carson, R. et Daigle, D. (2012). Une analyse des facteurs de risque dans l'apprentissage de la lecture chez l'enfant. *Enfance en difficulté, 1*, 47-83.
- Dowling, W.J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics, 14*(1), 37-40.
- Ecalte, J. (1997). Les représentations chez le lecteur novice ou expert : perspective componentielle. *Enfance, 285*-303.
- Ehri, L.C. (2005). Development of Sight Word Reading: Phases and Findings. Dans Snowling, M. J. et C. Hulme (dir.), *The science of reading: A handbook* (p. 135-154). Oxford, UK : Blackwell.
- Elman, J.L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences, 8*(7), 301-306.

- Engle, R.W., Cantor, J. et Carullo, J.J. (1992). Individual Differences in Working Memory and Comprehension: A Test of Four Hypotheses. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 18(5), 972-992.
- Engle, R.W. *et al.* (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology: General*, 128(3), 309-331.
- Fayol, M. (1992). *La compréhension lors de la lecture : un bilan provisoire et quelques questions*. Lille.
- Findley, L., Kohen, D. et Miller, A. (2014). Developmental milestones among Aboriginal children in Canada. *Paediatr Child Health*, 19(5), 241-246.
- Fortier, V. (2013). *Exploration de la relation entre les habiletés métasyntaxiques et la capacité de mémoire phonologique chez les enfants de langues d'origine*. (Thèse). Université du Québec à Montréal, Montréal.
- Fortin, L. et Picard, Y. (1999). Les élèves à risque de décrochage scolaire : facteurs discriminants entre décrocheurs et persévérants. *Revue des sciences de l'éducation*, 25(2), 359.
- Foucambert, D. (1994). *Analyse de textes*. Paris : Association française pour la lecture.
- Foucambert, D. (2003). *Syntaxe, vision parafovéale et processus de lecture : contribution du modèle structural à la pédagogie*. Université Grenoble II.
- Foucambert, D. (2007). *Quand la syntaxe guide la compréhension en lecture: premières approches issues du modèle structural de lecture*. 15th European Conference on Reading, Actes du colloque, 2007, Humboldt University, Berlin
- Freebody, P. et Anderson, R.C. (1983). Effects of Vocabulary Difficulty, Text Cohesion, and Schema Availability on Reading Comprehension. *Reading Research Quarterly*, 18(3), 277-294.
- French, L., Simard, D. et Fortier, V. (2012). L'utilisation d'une mesure de la mémoire phonologique auprès d'une population d'enfants linguistiquement hétérogène. *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 68(3), 291-315.

- Gathercole, S.E. (2009). Working memory and Language. Dans Gaskell, M. (dir.), *The Oxford handbook of psycholinguistics* : Oxford ; Toronto: Oxford University Press.
- Gathercole, S.E. et Baddeley, A.D. (1996). *The Children's Test of Nonword Repetition*. London, UK.
- Gernsbacher, M.A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. Dans Bower, G. H. (dir.), *The psychology of learning and motivation* (p. 217-263). New York : Academic Press.
- Gough, P.B. et Tunmer, W.E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, 7(1), 6-10.
- Hemphill, L. et Tivnan, T. (2008). The Importance of Early Vocabulary for Literacy Achievement in High-Poverty Schools. *Journal of Education for Students Placed at Risk (JESPAR)*, 13(4), 426-451.
- Hoover, W.A. et Gough, P.B. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- Hoover, W.A. et Tunmer, W.E. (1993). The Components of Reading. Dans Thompson, G. B., W. E. Tunmer et T. Nicholson (dir.), *Reading acquisition processes*. Clevedon, PA : Multilingual Matters.
- Howell, D.C. (2008). *Méthodes statistiques en sciences humaines, 2e édition*.
- Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, 21(1), 47-77.
- Johnson-Laird, P.N. (1983). *Mental Models*. Cambridge.
- Joshi, R.M. et Aaron, P.G. (2000). The Component Model of Reading: Simple View of Reading Made a Little More Complex. *Reading Psychology*, 21(2), 85-97.
- Just, M.A. et Carpenter, P.A. (1992). A Capacity Theory of Comprehension: Individual Differences in Working Memory. *Psychological Review*, 99(1), 122-149.
- Khomsy, A. (1999). *Lecture de mots et compréhension, révisée*. Paris : Éditions du Centre de psychologie appliquée.

- Kibby, M.Y. (2009). There are multiple contributors to the verbal short-term memory deficit in children with developmental reading disabilities. *Child Neuropsychol*, 15(5), 485-506.
- Klapp, S.T., Marshburn, E.A. et Lester, P.T. (1983). Short-term memory does not involve the "working memory" of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General*, 112(2), 240.
- Larson-Hall, J. (2012). *A guide to doing statistics in Second language research using SPSS*. NY: Routeledge.
- Lebart, L., Morineau, A. et Piron, M. (1995). *Statistique exploratoire multidimensionnelle*. Paris: Dunod.
- Lecocq, P. et al. (1996). *Apprentissage de la lecture et compréhension d'énoncés*. Paris : Presses Universitaires du Septentrion.
- Lervåg, A. et Aukrust, V.G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, 51(5), 612-620.
- Levelt, W.J.M. (1989). *Speaking: From Intention to Articulation*. Cambridge, M.A. : Bradford.
- Liberman, I.Y. et Shankweiler, D. (1989). Phonologie et apprentissage de la lecture : une introduction. Dans Rieben, L. et C. A. Perfetti (dir.), *L'Apprenti lecteur : recherches empiriques et implications pédagogiques* (p. 23-42) : Neuchâtel, Suisse Delachaux & Niestlé.
- Lurin, J. et Soussi, A. (2013). *Évolution des compétences en lecture et en écriture au cycle d'orientation chez des élèves à risque : étude longitudinale*. (Département de l'instruction publique). Genève : Service de la recherche en éducation.
- Magnusson, E. et Naclér, K. (1991). On the development of reading in good and poor readers. *International Journal of Applied Linguistics*, 1(2), 174-185.
- Maïonchi-Pino, N. (2008). *Le traitement syllabique chez l'enfant normo lecteur et dyslexique : rôle des caractéristiques linguistiques du français*. Université Lumière Lyon 2. Thèse de doctorat.

- Martinez Perez, T., Majerus, S. et Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: evidence from a longitudinal study. *Journal of Experimental Child Psychology*, 111(4), 708-723.
- McClelland, J.L. et Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of Basic Findings. *Psychological Review*, 88, 375-407.
- McCutchen, D. (1996). A Capacity Theory of Writing: Working Memory in Composition. *Educational Psychology Review*, 8(3), 299-325.
- Melby-Lervåg, M. (2011). The Relative Predictive Contribution and Causal Role of Phoneme Awareness, Rhyme Awareness and Verbal Short-Term Memory in Reading Skills: A Review. *Scandinavian Journal of Educational Research*, 56(4), 363-380.
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97
- Miller, G.A. (1996). *The Science of Words*. New York : Freeman.
- Morissette, C. (2012). *Pour se laver, l'ours brun utilise... une pierre!* Récupéré le 10 février 2013 de http://www.lesdebrouillards.qc.ca/client/page_article.asp?page=6306&clef=29&clef2=4
- Morris, L. (2013). The Effects of a Parent-child Reading Project on the Development of French Prereading Skills in Innu-speaking Kindergartners. *The Canadian Journal of Applied Linguistics*, 16(1), 135-151.
- Morris, L. (2014). *Rapport – École Nussim : Évaluation des élèves en innu-aimun et en français 2013-2014*.
- Morris, L. et Labelle, M. (2008). *Une juste mesure : Développement d'instruments et de critères d'évaluation linguistique pour des élèves allophones*.
- Morris, L. et Mackenzie, M. (2012). *Assessing the lexical knowledge of Innu-speaking children*. *Algonquian Conference*, Actes du colloque, 2012, Chicago
- Morris, L. et Mackenzie, M. (2013). Using all the Pieces to Solve the Puzzle: the Importance of Aboriginal Language Assessment in Child Populations. Dans Norris , M. J., D. Patrick et N. Ostler (dir.), *Endangered Languages Beyond*

- Boundaries: Community Connections, Collaborative Approaches and Cross-Disciplinary Research* (p. 170-177). Ottawa : Carleton University.
- Nagy, W.E. (2005). Why vocabulary instruction needs to be long-term and comprehensive. Dans Hiebert, E. H. et M. L. Kamil (dir.), *Teaching and learning vocabulary: Bringing research to practice* (p. 27-44). Mahwah : NJ: Lawrence Erlbaum.
- Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge : Cambridge University Press.
- Netten, A., Droop, M. et Verhoeven, L. (2011). Predictors of reading literacy for first and second language learners. *Read Writ*, 24(4), 413-425.
- Oakhill, J. (1988). *Becoming a skilled reader*. New York.
- OCDE. (2012). *Principaux résultats de l'Enquête PISA 2012 : Ce que les élèves de 15 ans savent et ce qu'ils peuvent faire avec ce qu'ils savent*. : Programme International pour le Suivi des Acquis des élèves.
- Oudin, A.-S. et Drapeau, L. (1993). Langue et identité ethnique dans une communauté montagnaise bilingue. *Revue québécoise de linguistique*, 22(2), 75.
- Ouellet, M.-C. (2012). *Vidéo : une bouteille lancée à la mer livre son message!*
Récupéré le 10 février 2013 de
http://www.lesdebrouillards.qc.ca/client/page_article.asp?page=6578
- Ouellet, M.-C. (2013). *Un jeune de 15 ans met au point un test de dépistage du cancer.*
Récupéré le 10 février 2013 de
http://www.lesdebrouillards.qc.ca/client/page_article.asp?page=6710&clef=&clef2=4
- Parrila, R., Kirby, J.R. et McQuarrie, L. (2004). Articulation Rate, Naming Speed, Verbal Short-Term Memory, and Phonological Awareness: Longitudinal Predictors of Early Reading Development? *Scientific Studies of Reading*, 8(1), 3-26.
- Perfetti, C.A. (1977). Language comprehension and fast decoding: Some psycholinguistic prerequisites for skilled reading comprehension. Dans Guthrie, J. (dir.), *Cognition, curriculum and comprehension*. Newark, DE : International Reading Association.

- Perfetti, C.A. (1985). *Reading ability*. New York : New York Oxford University Press.
- Perfetti, C.A. (1989). Représentations et prise de conscience au cours de l'apprentissage de la lecture. Dans Laurence Rieben, é. et é. Charles Perfetti (dir.), *L'Apprenti lecteur : recherches empiriques et implications pédagogiques* : Neuchâtel, Suisse Delachaux & Niestlé.
- Perfetti, C.A. et Lesgold, A.M. (1977). Discourse comprehension and sources of individual differences.
- Peterson, L.R. et Peterson, M.J. (1959). Short-term retention of individual verbal items. *Journal of experimental psychology*, 58(3), 193-198.
- Plaut, D.C. *et al.* (1996). Understanding Normal and Impaired Word Reading: Computational Principles in Quasi-Regular Domains. *Psychological Review*, 103(1), 56-115.
- Poncelet, M. et Van der Linden, M. (2003). L'évaluation du stock phonologique de la mémoire de travail : élaboration d'une épreuve de répétition de non-mots pour population francophone. *Revue de Neuropsychologie*, 13(3), 377-407.
- Qian, D.D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 56, 282-308.
- Qian, D.D. (2002). Investigating the Relationship Between Vocabulary Knowledge and Academic Performance: An Assessment Perspective. *Language Learning*, 52(3), 513-536.
- Qian, D.D. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, 21(1), 28-52.
- Rastle, K. (2009). Visual word recognition. Dans Gaskell, M. (dir.), *The Oxford Handbook of psycholinguistics*. Oxford ; Toronto: Oxford University Press.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, 19(2), 12-25.
- Rémond, M. et Quet, F. (1999). Apprendre à comprendre l'écrit. Psycholinguistique et métacognition : l'exemple du CM2. *Repères*, 19, 203 – 224.
- Ricketts, J., Nation, K. et Bishop, D.V.M. (2007). Vocabulary Is Important for Some, but Not All Reading Skills. *Scientific Studies of Reading*, 11(3), 235-257.

- Roberts, R. et Gibson, E. (2002). Individual Differences in Sentence Memory. *Journal of Psycholinguistic Research*, 31(6), 573-598.
- Rossi, J.-P. (2008). *Psychologie de la compréhension du langage*. Bruxelles : Bruxelles : De Boeck.
- Seigneuric, A. et al. (2000). Working memory resources and children's reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, 13, 81-103.
- Shiotsu, T. et Weir, C.J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing*, 24(1), 99-128.
- Simard, D., Foucambert, D. et Labelle, M. (2014). Examining the Contribution of Metasyntactic Ability to Reading Comprehension Among Native and Non-Native Speakers of French. *International Journal of Bilingualism*
- Snow, C.E., Burns, S.M. et Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC : National Academy Press.
- Spira, E.G., Bracken, S.S. et Fischel, J.E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, 41, 225-234.
- Stahl, S.A. (1999). *Vocabulary development*. Cambridge : MA: Brookline Books.
- Staphorsius, G. et Krom, R. (1998). *CITO Reading Comprehension Tests*. Arnhem, Pays-Bas.
- Swanson, H.L. et Berninger, V. (1995). The Role of Working Memory in Skilled and Less Skilled Readers' Comprehension. *Intelligence*, 21, 83-108.
- Swanson, H.L. et Berninger, V. (1996). Individual Differences in Children's Working Memory and Writing Skill. *Journal of Experimental Psychology: General*, 63, 358-385.
- Swanson, H.L. et Howell Ashbaker, M. (2000). Working Memory, Short-term Memory, Speech Rate, Word Recognition and Reading Comprehension in Learning Disabled Readers: Does the Executive System Have a Role? *Intelligence*, 28, 1-30.
- Swanson, H.L. et Siegel, L.S. (2001). Learning Disabilities as a Working Memory Deficit. *Issues in Education*, 7, 1-48.

- Taft, M. (1992). *Reading and the mental lexicon*. Hove, UK : Lawrence Erlbaum Associates Ltd.
- Treimen, R. et Kessler, B. (2009). Learning to read. Dans Gaskell, M. (dir.), *The Oxford handbook of Psycholinguistics* : Oxford ; Toronto: Oxford University Press.
- Tunmer, W.E. (1989). Conscience phonologique et acquisition de la langue écrite. Dans Rieben, L. et C. Perfetti (dir.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques*. Paris.
- Tunmer, W.E. et Chapman, J.W. (2012). The simple view of reading redux: vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, 45(5), 453-466.
- Van der Linden, M. *et al.* (1995). Adaptation en langue française du « Reading Span Test » de Daneman et Carpenter (1980). *L'année psychologique*, 459-482.
- Van Dijk, T.A. et Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York.
- Verhoeven, L. et van Leeuwe, J. (2008). Prediction of the development of reading comprehension: a longitudinal study. *Applied Cognitive Psychology*, 22(3), 407-423.
- Whitehurst, G.J. (1997). Language processes in context: Language learning in children reared in poverty. *Research on communication and language disorders: Contribution to theories of language development*, 233-266.
- Willis, C.S. et Gathercole, S.E. (2001). Phonological short-term memory contributions to sentence processing in young children. *Memory*, 9(4-6), 349-363.
- Wilson VanVoorhis, C.R. et Morgan, B.L. (2007). Understanding Power and Rules of Thumb for Determining Sample Sizes. *Tutorials in Quantitative Methods for Psychology*, 3(2), 43-50.