UNIVERSITÉ DU QUÉBEC À MONTRÉAL

MODÉLISATION DES CONTRIBUTIONS D'HABILETÉS COGNITIVES ET LINGUISTIQUES À LA COMPRÉHENSION EN LECTURE : ENJEUX MÉTHODOLOGIQUES ET DIDACTIQUES

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN LINGUISTIQUE

PAR
ROSALIE BOURDAGES

FÉVRIER 2024

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse 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.12-2023). 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

La réalisation de la présente thèse n'aurait pas été possible sans les personnes extrêmement talentueuses et généreuses que j'ai eu l'immense chance de rencontrer au fil de cette épopée académique, longue de plusieurs années et grande de deux continents.

Mes remerciements sentis et ma profonde gratitude vont d'abord à Roxane et Denis. Vous avez tous deux fait pour moi bien plus que ce qu'une direction de thèse implique. Votre engagement et votre expertise en sciences de l'éducation et en psycholinguistique, votre soutien indéfectible et vos conseils éclairés ont été cruciaux pour la bonne marche du processus délicat qu'a été la rédaction de cette thèse. Merci de m'avoir inspirée, puis guidée à faire ce travail. Vos relectures minutieuses, marquées par vos éclairages complémentaires et pertinents sur les objets traités, ont certainement contribué à améliorer le présent manuscrit. Je remercie également les institutions qui m'ont donné les moyens financiers, logistiques et académiques ayant permis ma formation doctorale : la HEP Vaud et l'UQAM.

Ensuite, je tiens à exprimer ma reconnaissance envers les membres du jury pour leur expertise et pour le temps passé à évaluer cette thèse.

J'adresse également des remerciements chaleureux à mes collègues du projet Grafe'maire. Nina, Jean-Paul, Véronique, Sandy, Matthieu, Martine, Anouk, Jean-François, Florence, Daniel, Serge, Christopher : merci de m'avoir partagé votre passion pour la didactique du français, merci pour nos échanges toujours riches et dynamiques. Votre savoir collectif a joué un rôle déterminant dans ma recherche.

Je souhaite aussi témoigner de ma profonde gratitude envers les enseignant·es (et leurs élèves) qui ont partagé leur temps et leur expertise avec moi tout au long de la collecte de données. En particulier, j'aimerais remercier Bastian Irmiger et Jérémy Gurba de l'école de Morges, Sophie Poza, Sophie Gfeller et Lucien Gfeller d'Echallens, May Jaccot et Isabelle Chenal à Assens, Loyse Noverraz de l'école de Gland, et Patricia Pradervand, Josiane Roberti, Anne-Claire Pittet, Léa Pouly et Laurence Vallon Vuithier de l'école d'Yverdon-les-Bains.

Je salue également Marie Côté, qui a implémenté les épreuves de conscience syntaxique dans *Psychopy*. Merci à Marine et Catherine pour leur empathie face aux

aléas inhérents aux analyses statistiques et leurs judicieux conseils permettant d'y remédier.

Mes remerciements vont également à mes collègues et amis pour leurs rires, leur soutien et leurs encouragements tellement précieux. Vivette, ACB, Vanessa, Vincent, Monnier, Kelly, Claire, Sonia, Violeta, Dédée, Jojo, Maude, l'autre Maude, Pam, Leslie, Jonathan, Eric, Jérem, Bobi Brown, Paco : merci!

Merci à ma famille d'ici et de là-bas d'avoir su me soutenir. Ma mère a relu l'entièreté de ce manuscrit et je l'en remercie. Seul un amour inconditionnel peut expliquer qu'un individu s'inflige pareil supplice.

J'ai adoré accomplir ce travail de thèse, et c'est grâce à vous tou·tes. Merci !

TABLE DES MATIÈRES

REMERCIEMENTS ii

LISTE DES FIGURES xi

LISTE DES TABLEAUX xii

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES xiv

RÉSUMÉ xv

ABSTRACT xvii

INTRODUCTION 1

PARTIE 1 CADRE THÉORIQUE 9

CHAPITRE I LES HABILETÉS MÉTALINGUISTIQUES 9

1.1 Métacognition 10

1.1.1 Remarques préliminaires 10
1.1.2 Développement de la métacognition 14
1.1.3 Taxonomie de la métacognition 17
1.1.4 Synthèse du cadre sur la métacognition 20

1.2 Définitions de la conscience métalinguistique et statut des construits métalinguistiques 21

1.2.1 Quelques consensus 24

1.2.1.1 « Métalinguistique » n'est pas « métalangage » 25
1.2.1.2 La langue comme objet 27
1.2.1.3 Le « métalinguistique » par opposition au « linguistique » 29
1.2.1.4 La notion de conscience 31

1.2.2 Dimensions sous-jacentes constitutives des habiletés métalinguistiques 35

1.2.2.1 Analyse des connaissances linguistiques 37
1.2.2.2 Contrôle cognitif 39
1.2.2.3 Représentations et processus 40

1.3 Synthèse du cadre sur les habiletés métalinguistiques 42

CHAPITRE II LA CONSCIENCE MORPHOLOGIQUE 45

2.1 Représentations morphologiques 46

2.1.1 Conscience morphologique 49
2.1.2 Développement de la conscience morphologique 50

2.2 La mesure de la conscience morphologique 54

2.2.1 Dimensionnalité de la conscience morphologique 55
2.2.2 Taxonomie des épreuves de conscience morphologique 57

2.2.2.1 Modalités de présentation 57
2.2.2.2 Types de morphèmes (unités) 58
2.2.2.3 Opérations 58
2.2.2.4 Analyse et niveaux de contrôle 59

2.3 Lien entre conscience morphologique et compréhension en lecture 62

2.3.1 Sensibilité au traitement sémantique 64
2.3.2 Calcul syntaxique 66

2.4 Synthèse du cadre théorique sur la conscience morphologique 66

CHAPITRE III LA CONSCIENCE SYNTAXIQUE 70

3.1 Représentations syntaxiques 71

3.1.1 Conscience syntaxique 76
3.1.2 Développement métasyntaxique 77

3.2 La mesure des habiletés métasyntaxiques 79

3.2.1 Éléments convergents 79
3.2.2 Description des épreuves métasyntaxiques 83
3.2.3 Dimensionnalité de la conscience syntaxique 86

3.2.3.1 Niveaux de contrôle 87
3.2.3.2 Analyse des connaissances syntaxiques 89
3.2.3.3 Types de syntagmes (unités syntaxiques) 91
3.2.3.4 Opérations 93

3.3 Lien entre habiletés métasyntaxiques et lecture 94

3.4 Synthèse du cadre théorique sur la conscience syntaxique 99

CHAPITRE IV LA COMPRÉHENSION EN LECTURE 101

4.1 La compréhension du discours : remarques préliminaires 102

4.2 Approches multicomponentielles de la compréhension 104

4.3 Niveaux de représentation 105

4.3.1 Structure de surface (*surface code*) 107
4.3.2 Base propositionnelle du texte (*textbase*) 108
4.3.3 Modèle de situation 111

4.4 Processus 114

4.4.1 Approches ascendante, descendante et interactive de l'acte de lire .. 115
4.4.2 Processus de bas niveau 119

4.4.2.1 Décodage 119
4.4.2.2 Accès au lexique 120

4.4.3 Processus de haut niveau 122

4.4.3.1 Construction, appariement et modification des modèles mentaux … 123
4.4.3.2 Activation et assemblage 124
4.4.3.3 Raisonnement déductif et inférences 125

4.5 Mesure de la compréhension en lecture 128

4.5.1 Perspectives diagnostique et explicative de la compréhension 128
4.5.2 Épreuves de compréhension dans les études sur la CM et la CS 131

4.6 Synthèse du cadre théorique sur la compréhension en lecture 135

CHAPITRE V SYNTHÈSE 137

PARTIE 2 CADRE MÉTHODOLOGIQUE 143

CHAPITRE VI INSTRUMENTS DE MESURE 143

6.1 Description des épreuves 144

6.2 Instruments de mesure 147

6.2.1 Épreuves de conscience morphologique 148

6.2.1.1 Paires de mots 150
6.2.1.2 Extraction de la base 152
6.2.1.3 Dérivation 153
6.2.1.4 Décomposition 155

6.2.1.5 Choix de pseudomots suffixés (morphèmes dérivationnels) 156
6.2.1.6 Choix de pseudomots suffixés (morphèmes flexionnels) 158
6.2.1.7 Portrait global des épreuves de conscience morphologique et dimensions impliquées dans leur traitement 159

6.2.2 Épreuves de conscience syntaxique 160

6.2.2.1 Déroulement 161
6.2.2.2 Éléments syntaxiques impliqués 164
6.2.2.3 Critères de construction des phrases 166
6.2.2.4 Contrebalancement 169
6.2.2.5 Ordre de difficulté dans les tâches 170
6.2.2.6 Limites des tâches de conscience syntaxique 171
6.2.2.7 Portrait global des épreuves de conscience syntaxique et dimensions impliquées dans leur traitement 173

6.2.3 Épreuves matricielles avec matériau non linguistique (*domain-general*) 174

6.2.3.1 Classifications 175
6.2.3.2 Matrices 176

6.2.4 Épreuves de compréhension en lecture 180

6.2.4.1 Épreuve 10 textes courts 182
6.2.4.2 Description de la variabilité dans les textes 188
6.2.4.3 Épreuve avec texte long 192

CHAPITRE VII DESIGN EXPÉRIMENTAL 194

7.1 Design expérimental 195

7.2 Prétests 196

7.3 Traitement des données 197

PARTIE 3 RÉSULTATS ET PERSPECTIVES 201

CHAPITRE VIII ANALYSES DESCRIPTIVES ET INFÉRENTIELLES DES DONNÉES 201

8.1 Rappel des questions de recherche 202

8.2 Préparation des données 202

8.2.1 Préparation des fichiers de données 202
8.2.2 Procédure 204
8.2.3 Description de l'échantillon 205

8.3 Statistiques descriptives pour les variables mesurées 207

8.4 Corrélations 213

8.5 Analyses en composantes principales..... 217

8.6 Spécifications des modèles AFC et SEM, et indices d'ajustement..... 221

8.6.1 Gestion des données manquantes..... 224
8.6.2 Méthode d'estimation 226
8.6.3 Indices d'ajustement et informations sur les modèles..... 226
8.6.4 Multicolinéarité 228
8.6.5 Comparaison de modèles 228

8.7 Plan d'analyse des modèles SEM et AFC appliquées aux variables de CM et de CS 229

8.8 Analyses appliquées aux variables de CM..... 231

8.8.1 Analyses factorielles confirmatoires appliquées à la CM 231

8.8.1.1 Modèle unidimensionnel de la CM 231
8.8.1.2 Modèle bidimensionnel de la CM impliquant les unités..... 234
8.8.1.3 Modèle bidimensionnel de la CM impliquant les opérations..... 235

8.8.2 Modélisations par équation structurelle (SEM) appliquées à la CM.... 237

8.8.2.1 Modèle unidimensionnel de la CM 238
8.8.2.2 Modèle bidimensionnel de la CM impliquant les opérations..... 240
8.8.2.3 Modèles bidimensionnels de la CM impliquant les unités 240

8.8.3 Synthèse des résultats sur la conscience morphologique 243

8.9 Analyses appliquées aux variables de CS..... 243

8.9.1 Analyses factorielles confirmatoires appliquées à la CS..... 243

8.9.1.1 Modèle unidimensionnel de la CS..... 244
8.9.1.2 Modèle bidimensionnel de la CS impliquant les opérations 244
8.9.1.3 Modèle bidimensionnel de la CS impliquant les unités 245

8.9.2 Modélisations par équations structurelles (SEM) appliquées à la CS 250

8.9.2.1 Modèle bidimensionnel CS impliquant les unités 250

8.9.3 Synthèse des résultats sur la conscience syntaxique 253

8.10 Résumé des résultats et limites de la recherche 253

CHAPITRE IX DISCUSSION 256

9.1 Réponses à la première question de recherche 257

9.2 Réponses à la deuxième question de recherche..... 260

9.3 Biais potentiels et pistes d'amélioration..... 265

9.4 Positionnement plus général sur le traitement de l'écrit 269

CHAPITRE X PERSPECTIVES DIDACTIQUES SUR LES TÂCHES MÉTASYNTAXIQUES 273

10.1 Présentation de la démarche de transposition didactique 274

10.2 Évaluation et enseignement de la grammaire 278

10.2.1 Enseignement de la grammaire 278
10.2.2 Activités syntaxiques et évaluation 280

10.3 Comparaison entre les tâches métasyntaxiques et les activités grammaticales 284

10.3.1 Enseignement de la grammaire 284
10.3.2 Activités syntaxiques et évaluation 287

10.4 L'erreur comme levier didactique 290

10.5 De l'erreur au méta, au raisonnement 291

10.6 Dispositif didactique proposé 296

10.7 Synthèse des perspectives didactiques en lien avec un dispositif basé sur l'erreur 300

CONCLUSION 305

ANNEXE A Protocole de passation, épreuves de conscience morphologique 309

ANNEXE B Protocole de passation, épreuves d'habiletés domain-general / non verbales 323

ANNEXE C Protocole de passation, épreuves de compréhension 326

ANNEXE D Protocole de passation, épreuves de conscience syntaxique 327

ANNEXE E Items des épreuves de conscience syntaxique 334

ANNEXE F Items des épreuves de conscience morphologique 339

ANNEXE G Items des deux épreuves de compréhension de textes 347

ANNEXE H Exemples d'items des épreuves matricielles avec matériau non linguistique 362

ANNEXE I Document incitatif destiné aux enseignant·es 363

ANNEXE J Formulaire de consentement pour les parents d'élèves 364

ANNEXE K Statistiques descriptives et corrélations pour les textes de l'épreuve avec textes courts 365

ANNEXE L Statistiques descriptives pour les épreuves mesurant les quatre grands construits, en fonction du niveau scolaire 371

ANNEXE M Épreuves et nom des variables associées 376

RÉFÉRENCES 378

LISTE DES FIGURES

Figure 1.1 Le modèle de Bialystok et Ryan (1985, p. 232), puis de Bialystok (2001a, p. 16), tiré de la traduction de Fortier (2013, p. 22)..... 38
Figure 2.1 Proposition de représentation graphique des tâches mesurant la conscience morphologique, en fonction de leurs demandes en termes d'analyse et de contrôle, telles que définies par Bialystok (2001a)..... 62
Figure 3.1 Représentation arborescente de la phrase « Basile aime les mets épicés »,où [P [SN Basile] [SV [V aime] [SN [dét[les]] [N[mets][SAdj[épicés]]]. 72
Figure 3.2 Le modèle de Bialystok (2001a) avec projection des tâches psycholinguistiques issues de Besse, A.-S. *et al.* (2010), Simard, D. *et al.* (2017), Cohen (2011) et Lesaux et Siegel (2003)..... 90
Figure 6.1 Exemple d'item de la tâche de classifications du CFT1-R (Weiß et Osterland, 2013). 176
Figure 6.2 Exemple d'item du *Raven's Colored Progressive Matrices*, tel que reproduit dans Carpenter *et al.* (1990) 178
Figure 6.3 Exemple d'item de la tâche de classifications du CFT1-R (Weiß et Osterland, 2013). 179
Figure 7.1 Représentations graphiques des modèles à tester, en fonction des dimensions hypothétisées dans les épreuves de conscience morphologique 199
Figure 7.2 Représentations graphiques des modèles à tester, en fonction des dimensions hypothétisées dans les épreuves de conscience syntaxique 200
Figure 8.1 AFC du modèle unidimensionnel de la CM. 233
Figure 8.2 Représentation graphique du modèle SEM avec les épreuves de conscience morphologique comme construit unidimensionnel 239
Figure 8.3 AFC du modèle bidimensionnel de la CS impliquant les unités. 247
Figure 8.4 Représentation graphique du modèle SEM avec les épreuves de conscience syntaxique comme construit bidimensionnel, divisé en fonction des unités mises en œuvre 251
Figure 10.1 Modèle de Bialystok (2001a) avec projection d'activités d'enseignement grammatical issues de Bulea Bronckart, Marmy Cusin & Panchout-Dubois (2017) et Lepoire Duc & Sautot (2010) 282

LISTE DES TABLEAUX

Tableau 4.1 Vue d'ensemble des tests de compréhension les plus couramment utilisés dans les travaux sur les liens entre habiletés métalinguistiques et lecture 132
Tableau 6.1 Vue synthétique des construits théoriques étudiés, des épreuves servant à les mesurer et des logiciels de la collecte des données 146
Tableau 6.2 Déroulement de la collecte 147
Tableau 6.3 Aperçu des tâches de conscience morphologique utilisées et de leur classement en fonction de deux dimensions 159
Tableau 6.4 Aperçu des caractéristiques des items dans les tâches de conscience syntaxique 166
Tableau 6.5 Aperçu des tâches de conscience syntaxique utilisées et de leur classement en fonction de deux dimensions. 173
Tableau 6.6 Aperçu des tâches métalinguistiques et transformationnelles prévues et de leur classement en fonction de deux dimensions hypothétisées 179
Tableau 6.7 Analyse lexicométrique : portrait de la fréquence du vocabulaire et de la longueur des phrases dans l'épreuve de compréhension de textes courts selon la base Lextutor (Cobb, 2019)..... 184
Tableau 6.8 Statistiques descriptives pour les différents petits textes..... 190
Tableau 8.1 Statistiques descriptives des variables brutes 209
Tableau 8.2 Statistiques descriptives des variables ayant subi une transformation 212
Tableau 8.3 Corrélations entre les variables à l'étude..... 215
Tableau 8.4 Contributions des facteurs de l'ACP avec valeur propre supérieure à 1 218
Tableau 8.5 Contributions des facteurs de l'ACP avec la méthode d'extraction basée sur une analyse en parallèle 220
Tableau 8.6 Indices d'ajustement pour les modèles des AFC appliquées aux épreuves de CM..... 236
Tableau 8.7 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles des AFC appliquées aux épreuves de CM 236
Tableau 8.8 Indices d'ajustement pour les modèles des SEM appliquées aux épreuves de CM..... 242
Tableau 8.9 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles des SEM appliquées aux épreuves de CM..... 242
Tableau 8.10 Indices d'ajustement pour les modèles AFC appliquées aux épreuves de CS 249

Tableau 8.11 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles AFC appliquées aux épreuves de CS 249
Tableau 8.12 Indices d'ajustement pour le modèle de SEM appliquée aux épreuves de CS 252
Tableau 8.13 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles des SEM appliquées aux épreuves de CS 252
Tableau 9.1 Aperçu des tâches de conscience morphologique utilisées et de leur classement en fonction de quatre dimensions 263
Tableau 9.2 Aperçu des tâches de conscience syntaxique utilisées et de leur classement en fonction de quatre dimensions 264
Tableau 10.1 Recension des tâches utilisées dans la recherche psycholinguistique pour mesurer les habiletés métasyntaxiques et des activités grammaticales réalisées en salle de classe 286
Tableau 10.2 Proposition d'activité grammaticale sur le complément du nom mettant en jeu la réplication de l'erreur 298

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

β	Coefficient bêta standardisé
$\Delta\chi 2$	Différence de valeurs de khi carré
$\chi 2$	Khi carré
AFC	Analyses factorielles confirmatoires (*Confirmatory Factor Analysis*)
AIC	*Akaike Information Criterion*
CD	Complément direct
CdN	Complément du nom
CI	Complément indirect
CFI	*Comparative Fit Index* (indice d'ajustement)
CM	Conscience morphologique
CS	Conscience syntaxique
Dét	Déterminant
GFI	*Goodness-of-Fit Index* (indice d'ajustement)
GAdj	Groupe adjectival
GN	Groupe nominal
Gprép	Groupe prépositionnel
GV	Groupe verbal
HmL	Habiletés métalinguistiques
L1	Langue première
L2	Langue seconde
PER	Plan d'études romand
Prép	Préposition
PSub	Phrase subordonnée
RMSEA	*Root-mean Square Error of Approximation* (indice d'ajustement)
SEM	Modélisation par équation structurelle (*Structural Equation Modelling*)
SRMR	*Standardized Root Mean Square Residual* (indice d'ajustement)
TLI	*Tucker-Lewis Index* (indice d'ajustement)
V	Verbe
VD	Variable dépendante
VI	Variable indépendante
ZPD	Zone proximale de développement

RÉSUMÉ

Un nombre grandissant de recherches en psycholinguistique démontrent l'importance de la conscience morphologique et de la conscience syntaxique en tant que prédicteurs de compétences linguistiques de haut niveau comme la compréhension en lecture (Gombert *et al.*, 1994; Kieffer *et al.*, 2016; Tong *et al.*, 2014). Ces habiletés métalinguistiques, réunies sous le vocable « conscience grammaticale », consistent en l'identification et la manipulation consciente des structures morphologiques et syntaxiques (Gombert, 1990; Pratt *et al.*, 1984), et permettent de discriminer les lecteur·trices expert·es des moins habiles (Nagy, 2007; Tong *et al.*, 2014). Gombert (2006) considère que les habiletés métalinguistiques évoluent sous l'effet de l'exposition à l'écrit, et l'apprentissage de la lecture lui-même est régulièrement décrit comme « fondamentalement métalinguistique » (Nagy et Anderson, 1999, p. 5).

La diversité des tâches utilisées pour mesurer ces habiletés (Goodwin *et al.*, 2017; Simard, D. *et al.*, 2017) a donné lieu à un débat méthodologique sur leur typologie et leur dimensionnalité présumée (Bialystok et Ryan, 1985; Goodwin *et al.*, 2017). Les épreuves élaborées pour mesurer la conscience morphologique et la conscience syntaxique semblent faire fonctionner deux types de variables. D'une part, les tâches se déclinent sur plusieurs unités de traitement – morphèmes dérivationnel ou flexionnel, mot ou syntagme. Mais réussir ces épreuves métalinguistiques nécessite d'autre part une habileté dans l'opération spécifique qui doit être réalisée, soit l'ajout et la suppression, indépendamment de leur contenu. Se dessinent alors deux grandes familles d'habiletés sous-jacentes de natures bien différentes : l'une est linguistique, tandis que l'autre est cognitive. La présente thèse cherche à construire des variables latentes représentant les habiletés transversales aux épreuves traditionnelles de conscience morphologique et de conscience syntaxique, et à évaluer expérimentalement l'influence de ces habiletés de natures linguistique et cognitive sur la réussite en compréhension en lecture.

Deux cent huit élèves de Suisse romande âgé·es de 8 à 12 ans ont été invité·es à effectuer trois tâches de conscience syntaxique (identification, correction et reproduction d'erreurs), six tâches de conscience morphologique (paires de mots, extraction de la base, dérivation, décomposition, choix de suffixes dérivationnels et flexionnels dans des pseudomots) et deux tâches transformationnelles non linguistiques. La compréhension en lecture a été mesurée à l'aide de deux tâches.

La dimensionnalité des tâches morphologiques et syntaxiques a été d'abord explorée grâce à des analyses factorielles confirmatoires (AFC). À la suite de cette analyse, l'influence des habiletés latentes construites sur la compréhension en lecture a été examinée à l'aide de modèles par équation structurelle (SEM). Les résultats suggèrent que l'habileté à prendre conscience de l'unité linguistique en jeu dans les épreuves est déterminante pour la compréhension en lecture. Les opérations minimales de nature cognitive, telles que nous les avons conçues, ne jouent pas un rôle significatif dans les mécanismes en jeu dans les habiletés métalinguistiques.

Le projet cherche également à mieux comprendre le statut des habiletés métalinguistiques au sein de la didactique, en le mettant en parallèle avec le statut de ces mêmes habiletés au sein de la psycholinguistique. La thèse cherche à ouvrir le dialogue sur des pistes de transposition didactique, en prenant position sur les types d'activités qui seraient favorables au raisonnement grammatical, en passant de la mesure à l'évaluation.

Mot-clés : psycholinguistique, mesures langagières et leurs modélisations statistiques, habiletés métalinguistiques, compréhension en lecture, psychologie appliquée à l'éducation, didactique.

ABSTRACT

English title : MODELING THE CONTRIBUTIONS OF COGNITIVE AND LINGUISTIC SKILLS TO READING COMPREHENSION: METHODOLOGICAL AND DIDACTIC ISSUES

A growing body of psycholinguistics research is demonstrating the importance of morphological awareness and syntactic awareness as predictors of high-level language skills such as reading comprehension (Gombert *et al.*, 1994; Kieffer *et al.*, 2016; Tong *et al.*, 2014). These metalinguistic skills, also referred to as “grammatical awareness”, consist of the identification and conscious manipulation of morphological and syntactic structures (Gombert, 1990; Pratt *et al.*, 1984), and make it possible to discriminate expert readers from the less skilled (Nagy, 2007; Tong *et al.*, 2014). Gombert (2006) considers that metalinguistic skills evolve with exposure to printed words, and learning to read itself is often described as “fundamentally metalinguistic” (Nagy and Anderson, 1999, p. 5).

The diversity of tasks used to measure these skills (Goodwin *et al.*, 2017; Simard, D. *et al.*, 2017) has given rise to a methodological debate about their typology and presumed dimensionality (Bialystok and Ryan, 1985; Goodwin *et al.*, 2017). The tests developed to measure morphological awareness and syntactic awareness appear to involve two types of variables. On the one hand, the tasks span several processing units – derivational or inflectional morphemes, words or phrase. On the other hand, successful completion of these metalinguistic tests requires skill in the specific operation to be carried out, i.e. addition and deletion, irrespective of their content. Two major families of underlying skills of quite different natures thus emerge: one is linguistic, while the other is cognitive. The aim of this thesis is to build latent variables representing skills crosscutting the traditional morphological and syntactic awareness tests, and to evaluate experimentally the influence of these linguistic and cognitive skills on reading comprehension.

Two hundred and eight students of 8 to 12 years old from French-speaking Switzerland were asked to complete three syntactic awareness tasks (identification, correction, and replication of errors), six morphological awareness tasks (“comes from” task, base extraction, derivation, decomposition, derivational suffix choice in non-words and inflectional suffix choice in non-words) and two non-linguistic transformational tasks. Reading comprehension was measured using two tasks.

The dimensionality of the morphological and syntactic tasks was first explored using confirmatory factor analysis (CFA). Following this analysis, the influence of these latent skills on reading performance will be examined, allowing for the disentangling of the relations between these constructs, using structural equation modelling (SEM). The results suggest that the ability to become aware of the linguistic unit at stake in metalinguistic tests is decisive for reading comprehension. Minimal cognitive operations appear to not play a significant role in the mechanisms involved in metalinguistic skills.

The project also seeks to better understand the status of metalinguistic skills within the field of didactics, by drawing a parallel between the status of these skills within psycholinguistics. This dissertation seeks to open up a dialogue on avenues for didactic transposition, taking a stance on the types of activities that are conducive to grammatical reasoning, moving from measurement to evaluation.

Keywords : Psycholinguistics, linguistic measurement, methods, SEM, model, metalinguistic awareness, reading comprehension, psychology applied to education, didactics.

INTRODUCTION

La maitrise[1] de la langue écrite, outre son rôle démocratique important à l'échelle sociale et communautaire, est garante de la réussite scolaire, de l'intégration sociale, de l'épanouissement personnel et du succès professionnel (Hemphill et Tivnan, 2008; Swanson *et al.*, 2013). La lecture, plus particulièrement, est travaillée de façon intensive à l'école primaire, et constitue ainsi le socle des apprentissages dans les autres disciplines scolaires (Berger, M. J. et Desrochers, 2011). Son développement revêt ainsi un intérêt particulier, tant pour les chercheur·es[2] en psycholinguistique et en acquisition que pour les didacticien·nes. La littérature scientifique dans ces domaines souligne depuis longtemps que les performances en lecture sont liées à l'analyse de la langue et à la prise en compte de son caractère structuré – en psycholinguistique : Demont et Gombert (1996); Gleitman et Rozin (1977); en didactique : Nadeau et Fisher (2006).

La manipulation de l'écrit est en effet une activité requérant le développement de connaissances et de stratégies spécifiques à la langue écrite, qui sont en partie

[1] Cette thèse applique les rectifications orthographiques adoptées par l'Académie française en 1990.
[2] Cette thèse adopte une forme d'écriture inclusive recommandée par l'Instance pour la promotion de l'égalité de la HEP Vaud.

différentes de celles nécessaires pour l'oral ou pour la conversation (Bialystok, 1988a; Gombert, 1991; Gombert *et al.*, 1994; Nagy, 2007; Ravid et Tolchinsky, 2002). Ces stratégies reposent sur l'adoption d'une position métalinguistique vis-à-vis du discours, c'est-à-dire sur l'attention qui sera portée au langage utilisé pour transmettre le sens, et pas seulement sur le sens lui-même (Le Ny, 2017; Nagy *et al.*, 2003). La compréhension en lecture est ainsi conçue comme une activité fondamentalement métalinguistique (Deacon et Kieffer, 2018).

Gombert (2006) considère que les habiletés métalinguistiques (désormais HmL), définies comme les habiletés à analyser et à manipuler les éléments constitutifs du langage, évoluent sous l'effet de l'exposition à l'écrit (Gombert, 1991). La mise en place et le développement des habiletés *linguistiques* n'impliquent pas un développement systématique ni une maitrise assurée des habiletés *métalinguistiques*. Celles-ci se déclinent sur au moins trois unités de traitement s'étant vu attribuer un rôle significatif dans la performance générale en lecture : phonologique (Chiape et Siegel, 1999; Demont et Gombert, 1996; Desrochers *et al.*, 2009; Nagy *et al.*, 2003), morphologique (Carlisle, 2000; Casalis *et al.*, 2009; To *et al.*, 2016) et syntaxique (Deacon et Kieffer, 2018; Foucambert, 2008; Gaux et Demont, 1997), en permettant de discriminer les lectrices et lecteurs expert·es des mauvais·es lectrices et lecteurs, tant en langue première (Gombert, 1990) que chez les apprenant·es d'une langue seconde (Besse, A.-S. *et al.*, 2010).

La compréhension de l'écrit est un construit[3] éminemment complexe qui s'appuie sur de nombreuses habiletés (Nagy, 2007). Un grand nombre de modèles multicomponentiels de la compréhension, mettant l'accent sur les composantes linguistiques et cognitives à sa source, ont comme ambition d'expliquer les différences interindividuelles importantes observées en compréhension en lecture (Hannon et Daneman, 2001). C'est dans ce contexte de prédiction des performances en lecture que les habiletés métalinguistiques ont acquis une place importante tant au sein de la recherche que dans les pratiques enseignantes.

Les travaux princeps sur le rôle des habiletés métalinguistiques dans l'apprentissage de la lecture se sont focalisés sur la conscience phonologique. Les lectrices et lecteurs débutant·es doivent être capables d'analyser la structure interne des mots pour automatiser la correspondance entre phonèmes et graphèmes. Le caractère prédictif de la conscience phonologique, très abondamment documenté (Bianco *et al.*, 2012; Gabriel et Poncelet, 2009; Melby-Lervåg, 2011; Sawyer et Fox, 1991), s'amenuise toutefois au fur et à mesure que s'affirme le décodage et que se déploient les stratégies de compréhension en lecture. Sont cependant documentés des groupes d'enfants aux prises avec d'importants déficits sur le plan de la compréhension, spécifiquement, et ce, malgré des habiletés normales en conscience phonologique, en lecture de mots et en vocabulaire (Cain et Oakhill, 2007; Tong *et al.*, 2014). Ainsi,

[3] Le terme « construit », ou *construct* en anglais, désigne dans le cadre de cette thèse une notion correspondant au résultat d'une construction théorique et modélisable. Il fait écho à son domaine princeps d'appartenance, à notre connaissance : la psychologie.

l'intérêt suscité par les deux construits représentant la « conscience grammaticale »[4] (Blakmore et Pratt, 1997; Bowey, 1994), soit la conscience morphologique et la conscience syntaxique, est en grande partie attribuable à leur faculté de prédire la compréhension en lecture chez des enfants d'âge scolaire (Carlisle, 2000; Deacon et Kirby, 2004; Demont et Gombert, 1996; Nagy *et al.*, 2006; Tong *et al.*, 2014). La littérature sur le caractère causal de ces liens, avec médiation par d'autres habiletés (Li et Wu, 2015) ou non (Tong *et al.*, 2014), est foisonnante. Les *patterns* de performance des « mauvais compreneurs » et « mauvaises compreneuses » sont assez mitigés, et laissent penser que les faiblesses en conscience morphologique et en conscience syntaxique sont tributaires des caractéristiques particulières des tâches permettant de les mesurer (Tong *et al.*, 2014). Cet état de choses suggère qu'il faut porter un regard pointu sur la mesure de ces construits dans la littérature en psychologie cognitive, afin de mieux comprendre ce qui explique les variations dans les performances en compréhension.

Plus récemment, les HmL ont donc fait l'objet de discussions méthodologiques, notamment dans le but de vérifier ce que mesurent réellement les épreuves qui sont censées les représenter (*construct validity*). La disparité des instruments de mesure utilisés dans la littérature sur les HmL a fait émerger une série d'enjeux ayant trait à la pratique de leur évaluation et aux habiletés sous-jacentes, ou dimensions, qu'elles mettent en œuvre (Berthiaume *et al.*, 2010; Bourdages et Foucambert, 2018; Carroll

[4] Le concept de *grammatical awareness* est plus usuel dans la littérature psycholinguistique anglophone.

et al., 2003; Kieffer *et al.*, 2016; Schatschneider *et al.*, 1999). Les épreuves élaborées pour mesurer les HmL font fonctionner deux types de variables qui se situent chacune sur un continuum : d'une part, le type d'unité traitée – phonème, morphème, groupe syntaxique, etc. – et, d'autre part, l'opération demandée – élision, ajout, inversion, remplacement (Bernstein *et al.*, 2020; Demont *et al.*, 2006; Desrochers *et al.*, 2009). Ainsi, réussir aux épreuves mesurant les différentes HmL nécessite également une habileté dans l'opération spécifique qui doit être réalisée, soit la suppression ou l'ajout – l'inversion ou le remplacement étant les produits de ces opérations – indépendamment de son contenu (Bernstein *et al.*, 2020; Tong *et al.*, 2014). Se dessinent alors deux grandes familles d'habiletés latentes de nature bien différente : l'une est linguistique, tandis que l'autre est cognitive. À notre connaissance, peu de travaux sont dédiés à l'étude systématique de la contribution respective de ces habiletés aux variables mesurées dans les épreuves traditionnelles d'HmL et l'influence de celles-ci sur la réussite en lecture.

La présente thèse a comme premier objectif de vérifier expérimentalement, à l'aide de la modélisation par équation structurelle (désormais SEM), si les mesures traditionnelles des habiletés en conscience morphologique et en conscience syntaxique sont subsumées dans deux grands ensembles distincts d'habiletés, le premier linguistique, représentant le type d'unités traitées, et le second cognitif, représentant le type d'opérations mentales effectuées. Notre première question de recherche, associée à ce premier objectif, est donc la suivante :

- *Les mesures traditionnelles des habiletés en conscience morphologique et en conscience syntaxique peuvent-elles se scinder en deux constituants, l'un linguistique, l'autre cognitif, prenant la forme de variables latentes ?*

Ces variables latentes seront mises en relation avec les performances générales en compréhension en lecture, ce qui nous amène à poser notre deuxième question de recherche :

- *Dans quelle mesure les variables latentes construites – d'ordres linguistique et cognitif – contribuent-elles à la compréhension en lecture ?*

Le projet cherche également à mieux comprendre le statut des habiletés « méta » au sein de la didactique, en le mettant en parallèle avec le statut de ces mêmes habiletés au sein de la psycholinguistique. L'hypothèse du lien entre habiletés métalinguistiques et compréhension est optimiste, puisqu'elle sous-tend que le progrès sur le plan de la compréhension peut se faire sur la base d'un travail sur des habiletés constitutives de la lecture qui peuvent être exercées et enseignées (Nagy, 2007). À travers une revue des différences et des ponts épistémologiques qui caractérisent l'étude du construit « méta » au sein des deux disciplines, la thèse cherche à ouvrir le dialogue sur des pistes de transposition didactique, par un passage de l'épreuve psycholinguistique à l'exercice en classe, de la mesure à l'exercisation, de l'épilinguistique à l'activité consciente. Ultimement, la thèse cherche à penser une forme d'autonomie de la

didactique, à travers une interaction avec la psychologie cognitive, en creusant le rapport de ces deux disciplines vis-à-vis du construit métalinguistique. Notre objectif est de comprendre comment peuvent se travailler les habiletés métalinguistiques en classe. Plus précisément, la question de recherche associée à ce troisième objectif est la suivante :

- *Comment peut-on opérer le passage entre tâches métalinguistiques et activités grammaticales?*

Ainsi, cette thèse, dont les enjeux sont d'ordres heuristique et théorico-méthodologique, trouve sa pertinence sociale à travers un dialogue entre la psycholinguistique et la didactique du français.

La thèse est organisée en trois parties. Dans la première partie, qui est constituée de cinq chapitres, nous jetons les bases du cadre théorique auquel nous ferons appel en présentant tout à tour les revues théorique et méthodologique des quatre construits qui nous intéressent (Chapitre I : Les habiletés métalinguistiques, Chapitre II : La conscience morphologique, Chapitre III : La conscience syntaxique, Chapitre IV : La compréhension en lecture, Chapitre V : Synthèse). La seconde partie de la thèse est dédiée au cadre méthodologique, le chapitre VI étant consacré à la présentation des instruments de mesure et le chapitre VII au design expérimental de notre étude. Au sein de la troisième et dernière partie, nous nous employons à présenter les résultats

et les perspectives psycholinguistiques et didactiques qu'ils laissent entrevoir (Chapitre VIII : Analyses descriptives et inférentielles, Chapitre IX : Discussion, Chapitre X : Perspectives didactiques sur les tâches métasyntaxiques). Enfin, la Conclusion fait état des principaux apports de la thèse.

PARTIE 1 CADRE THÉORIQUE

Knowledge rooted in experience shapes
what we value and as a consequence
how we know what we know as well as
how we use what we know.

— bell hooks, *Teaching Critical Thinking:*
Practical Wisdom

CHAPITRE I

LES HABILETÉS MÉTALINGUISTIQUES

L'intérêt de l'étude des habiletés métalinguistiques à travers le prisme des épreuves qui les mesurent a été brièvement démontré dans l'introduction. Nous nous intéresserons, dans ce chapitre, aux modèles théoriques qui président à l'élaboration des différentes épreuves métalinguistiques. Nous passerons en revue certains problèmes terminologiques apparaissant dans les travaux sur les habiletés métalinguistiques et les construits dits « méta » en général. Ainsi, dans cet effort de désambiguïsation, nous serons amenée à définir la métacognition (1.1) et les habiletés métalinguistiques (1.2), à travers ses caractéristiques convergentes (1.2.1) et ses dimensions sous-jacentes (1.2.2). Nous terminerons par une synthèse sur la métacognition telle que nous l'abordons dans ce travail (1.3).

1.1 Métacognition

Les activités métalinguistiques constituent pour Gombert (1990, p. 27) « un sous-domaine de la métacognition, qui concerne le langage et son utilisation ». La métacognition est définie par Flavell (1985, p. 104) comme la « cognition sur la cognition »[5], ou la conscience qu'un individu possède de ses propres processus cognitifs.

1.1.1 Remarques préliminaires

Les bases théoriques de l'étude contemporaine de la métacognition proviennent d'une part de la psychologie cognitive et de la psychologie du développement cognitif (Pintrich, 2002) ; elles sont représentées d'abord par les travaux de Piaget (1947), puis de Flavell (1976) et de Brown, A. L. (1987). Elles émergent, d'autre part, de la psychologie du développement social, telle que représentée par Vygotsky (Akturk et Sahin, 2011; Fox et Riconscente, 2008).

On entrevoit la métacognition, et en particulier l'importance de la prise de conscience dans la conceptualisation et la compréhension, dans l'avant-propos de « Réussir et comprendre » de Piaget (1974, p. 6)[6] :

---

[5] « *It is called* meta*cognition because its core meaning is "cognition about cognition". »* (Flavell, 1985, p. 104)

[6] La théorie piagétienne du développement prédit que le développement cognitif de l'individu se produit simultanément dans différents domaines de la connaissance (langage, pensée mathématique, etc.) Cette position a été contestée dans les années 80 par les positions sur la modularité de l'esprit, dont le principal porte-étendard est Fodor (1983). Les stades

Une première étape de notre analyse a fait l'objet d'un ouvrage préliminaire sur « La prise de conscience » de l'action propre et a consisté à montrer qu'il existe des actions complexes quoique de réussite précoce, et présentant tous les caractères d'un savoir, mais seulement d'un « savoir-faire », et que le passage de cette forme pratique de connaissances à la pensée s'effectuait au moyen de prise de conscience, mais ne se réduisant nullement à une sorte d'éclairage et consistant en une conceptualisation proprement dite, c'est-à-dire à une **transformation des schèmes d'action en notions et en opérations** : or cette modification fondamentale peut ne se produire que des années après la réussite pratique, la prise de conscience étant alors retardée par des déformations variées, allant jusqu'à des sortes de « refoulements » spectaculaires, le sujet ne parvenant pas à « voir » en ses propres actions certains caractères pourtant entièrement « observables » qui assurent leur succès, mais dont l'inconscience ou **l'absence d'enregistrement par la conscience empêche la compréhension conceptualisée**.

Dans le présent ouvrage il s'agissait d'abord de vérifier que cette autonomie et le caractère cognitif de l'action se maintiennent, antérieurement à la prise de conscience, jusque dans le cas des actions à réussite non précoce, mais s'effectuant par étapes et au moyen de coordinations de plus en plus complexes. Il s'agissait ensuite d'étudier le renversement progressif de cette situation, **lorsque la conceptualisation rejoint le niveau de l'action et finit, vers 11-12 ans, par le dépasser et par influencer en retour les actions jusqu'à pouvoir les diriger en les programmant avant toute réalisation**. Il s'agissait surtout, et c'est là notre but principal, de préciser les analogies et différences entre **« réussir », ce qui est la sanction du**

---

d'opérations formelles de Piaget s'appuient sur le déploiement de cette habileté à prendre pour objet ses propres processus cognitifs (Doudin et Martin, 1999). Les notions piagétiennes d'assimilation et d'accommodation, dans le modèle du développement cognitif chez l'enfant, apportent un cadre de référence général sur l'interaction entre les systèmes cognitifs et l'environnement, et sont des outils pertinents pour penser la maturation cognitive de l'individu. Pour Piaget (1947), la cognition humaine est la réponse biologique d'un système complexe pour appréhender un environnement non moins complexe (tel que mentionné par Flavell (1985)). Dans la théorie piagétienne sur le développement cognitif, *domain-general* dans son essence, le système cognitif humain est actif, et les processus d'assimilation et d'accommodation fonctionnent en interdépendance, tous deux mobilisés lorsqu'une information doit être traitée afin de permettre la construction de connaissances.

> **« savoir faire » et « comprendre », ce qui est le propre de la conceptualisation**[7], qu'elle succède à l'action ou finisse par la précéder et l'orienter. (Nous avons fait ressortir les éléments les plus pertinents en les mettant en gras.)

Cet extrait nous semble intéressant en ce qu'il exprime la thèse qui sous-tend toute l'entreprise métacognitive : c'est la prise de conscience qui permet la « compréhension conceptualisée » et virtuellement toute forme d'apprentissage, vers un dépassement et vers le pouvoir de diriger son attention et ses actions en fonction de la nature de la réalisation.

Le concept de métacognition est moins sujet à polémique sur le plan de sa définition que le sont les habiletés métalinguistiques (Akturk et Sahin, 2011; Garner et Alexander, 1989; Gombert, 1990), qui feront l'objet de la partie 1.2, et ce, malgré qu'il soit assez difficile à opérationnaliser sur le plan méthodologique (Woll, 2019). Il y a métacognition lorsque le système cognitif est engagé à réfléchir sur lui-même :

> « […] le consensus est total pour affirmer que la métacognition englobe les connaissances introspectives sur les états cognitifs et leurs opérations, et les capacités de l'individu de contrôler et de planifier ses propres processus de pensée et ses produits » (Gombert, 1990, p. 17).

[7] Les notions de « savoir-faire » et la « conceptualisation » dans cet extrait nous semblent pertinentes dans la mesure où elles évoquent d'une part, une part procédurale, un processus, et d'autre part, une unité conceptuelle, de nature déclarative, une représentation. Ces deux types d'habiletés, l'une de nature cognitive, l'autre plutôt de nature linguistique, jalonneront notre réflexion dans le cadre de la présente thèse.

La métacognition fait référence aux connaissances qui permettent de réguler, de contrôler et d'orchestrer les informations relatives à l'accomplissement d'une tâche cognitive (Flavell, 1976, 1985). La régulation, ou *monitoring* (Baker, 1989; Flavell, 1976, 1979, 1985; Son et Schwartz, 2002), se réfère à la capacité de poser un jugement sur la qualité de ses processus cognitifs[8], tandis que le contrôle permet d'utiliser ces éléments d'évaluation afin de modifier son comportement (Son et Schwartz, 2002).

Cette habileté à prendre conscience de la structure de ses propres processus de pensée et à la manière de les diriger et de les contrôler apparait comme essentielle à l'apprentissage, en particulier dans des domaines impliquant la résolution de problèmes et un matériel conceptuel difficile, comme la grammaire ou les mathématiques, et virtuellement dans tous les apprentissages scolaires. Les habiletés métacognitives sont réputées être constamment sollicitées dans la compréhension en lecture et le processus d'écriture, notamment (Flavell, 1985). La formation de concepts, qui résulte de la manipulation de concepts à l'aide du langage, est sous-tendue par la connaissance métacognitive. Une autre forme de connaissance métacognitive, mentionnée par Vygotsky (1930/1933/1978, 1934/1997), est la connaissance de ses propres capacités mentales, en particulier considérée par

[8] Le modèle de la mémoire d'Atkinson et Shiffrin (1968) constitue un revirement dans la littérature sur la mémoire, puisqu'il inclut des processus de conscientisation et de contrôle (Doudin et Martin, 1999). Ces processus sont susceptibles de varier en fonction des individus, à l'instar des capacités mnésiques, mais ils ne sont pas permanents ni intrinsèques à la mémoire : ils sont plutôt relatifs à la tâche demandée, à l'histoire du sujet et à son expérience à réussir des épreuves en adoptant des stratégies efficaces.

rapport à une tâche donnée. Les adultes ont cette connaissance, cette capacité d'abstraction réflexive, à des niveaux certainement divers, et les enfants la développent.

1.1.2 Développement de la métacognition

La progression marquée des habiletés métacognitives avec l'âge et l'expérience, notamment scolaire, laisse sous-entendre que la cognition, ou l' « intelligence », pour Vygotsky, est fonction du perfectionnement graduel des mécanismes permettant l'acquisition d'information, en permettant un contrôle et un guidage de plus en plus efficaces de l'activité en cours (Doudin et Martin, 1999). La métacognition est conceptuellement proche de l'autorégulation, conçue comme des phénomènes de contrôle et de révision des productions du sujet (Doudin et Martin, 1999). L'autorégulation requiert une prise de conscience intentionnelle de ses propres processus cognitifs (Fox et Riconscente, 2008). On retrouve les caractéristiques de l'autorégulation dans les écrits de Vygotsky (1934/1997), notamment dans ceux portant sur le concept de la zone proximale de développement (désormais ZPD). En effet, la ZPD est conçue comme un passage entre des mécanismes hétérorégulateurs mobilisés chez une personne tierce, comme un parent ou un·e enseignant·e et des mécanismes autorégulateurs, présents chez un sujet apprenant. Dans leur acception vygotskienne, ces phénomènes se développent avec l'âge et l'expérience, sont socialement construits, et sont essentiellement langagiers (Fox et Riconscente, 2008).

La métacognition se développe avec le temps (Flavell, 1985) et se montre sensible à l'expérience scolaire, lors de laquelle les élèves sont amené·es à faire des activités cognitives un objet d'analyse à part entière, applicable à des connaissances de natures différentes (Garner et Alexander, 1989). Les activités de résolution de problèmes chez les jeunes enfants sont marquées par un contrôle « épiprocédural » ou « épicognitif », l'attention consciente n'étant portée que sur la réussite ou l'échec de l'atteinte du but (Gombert et Fayol, 1988). On constate aussi un emploi réduit de stratégies de réparation de la compréhension (Baker et Brown, 1984) et une absence d'anticipation des processus (Gombert et Fayol, 1988). Les enfants plus âgé·es déploient des activités de planification et d'attention (Gombert et Fayol, 1988), par exemple, et reconnaissent l'importance de la concentration dans la réussite d'une tâche (Garner et Alexander, 1989).

Le développement de la métacognition chez l'enfant est aussi fonction d'autres capacités, comme l'attention ou la mémoire de travail. L'émergence d'une « attitude métacognitive » (Gombert et Fayol, 1988) est marquée par l'application de savoirs procéduraux et de procédures de résolution de problèmes assez coûteux cognitivement (Gombert et Fayol, 1988; Karmiloff-Smith, 1986). « Les modifications représentationnelles ou procédurales qui s'ensuivent sont contraintes par les certitudes "théoriques" des sujets. Le progrès cognitif se traduira donc ici souvent par une régression des performances » (Gombert et Fayol, 1988, p. 57).

L'application de schémas de résolution et de stratégies à des problèmes divers induit une diminution de la charge cognitive. L'évolution des habiletés métacognitives culmine par l'automatisation des processus (Gombert et Fayol, 1988). Cette automatisation par instanciation de schémas disponibles « n'est ni générale, ni complète, ni toujours stable » :

> Elle n'est pas générale dans la mesure où le sujet sera toujours susceptible de rencontrer de nouveaux problèmes non assimilables aux schémas dont il dispose en MLT. Elle n'est pas complète dans la mesure où il peut rencontrer des problèmes complexes pour lesquels il ne dispose pas de schéma d'ensemble, mais de schémas s'appliquant à tel ou tel sous-problème particulier (par exemple, une multiplication dans un problème arithmétique complexe). Elle n'est pas toujours stable dans la mesure où […] le sujet peut être conduit à assimiler un nouveau problème à un type de problème qu'il sait résoudre, il instanciera alors un schéma inapproprié, le comparateur final pouvant éventuellement lui permettre de détecter l'erreur et le conduire à reprendre la résolution avec une attitude métacognitive analytique » (Gombert et Fayol, 1988, p. 57).

Comme toute forme de connaissance, les connaissances métacognitives peuvent être erronées, ou manquer de nuances, et ultimement induire des ratés sur le plan métacognitif. En effet, les échecs de planification ou de contrôle des activités métacognitives, qui se manifestent par exemple par des lacunes dans la reconnaissance d'un bris de compréhension, ou d'incapacité à résoudre un problème logico-mathématique, se manifestent aussi à l'âge adulte (Baker, 1989). Il a été démontré que les personnes qui parviennent à résoudre des problèmes de manière efficace et créative passent un temps considérable à comprendre un problème avant

de mettre en œuvre des solutions, peut-être même plus que chez les personnes qui résolvent les problèmes de manière moins efficace, parce qu'elles prennent moins de temps pour les essais-erreurs (Berardi-Coletta *et al.*, 1995). Cet état de choses rend compte de l'importance de la pause, de la réflexion, dans la résolution de problèmes en général.

1.1.3 Taxonomie de la métacognition

Flavell (1976, 1979, 1985, 1987) a mis en place une taxonomie de la métacognition basée sur les concepts plus englobants de *connaissance* et d'*expérience* métacognitives[9]. Les *connaissances* métacognitives se réfèrent à l'ensemble des connaissances acquises par un individu sur les états et les processus psychologiques et cognitifs des autres et de lui-même[10], ou encore des connaissances relatives au contrôle de ces états et processus (Flavell, 1987). Selon l'auteur, celles-ci s'acquièrent de manière involontaire et automatique, et se subdivisent en trois catégories de variables distinctes : celles relatives aux *personnes*, celles relatives aux *tâches*, et celles relatives aux *stratégies*.

---

[9] Flavell (1978) définit la métacognition comme « la connaissance qui prend pour objet ou régule tout aspect de tout effort cognitif » (notre traduction de "*knowledge that takes as its object or regulates any aspect of any cognitive endeavor*"). Baker et Brown (1984) ont relevé le fait que cette définition inclut alors deux types de mécanismes ou d'habiletés : les connaissances sur la cognition et la régulation de la cognition.

[10] Le concept de « connaissances métacognitives » se rapproche de la théorie de l'esprit (*theory of mind*), qui désigne, en sciences cognitives, un ensemble d'habiletés permettant de s'attribuer à soi et aux autres des états mentaux (états affectifs et cognitifs, croyances, etc.).

Les connaissances métacognitives se référant aux personnes regroupent des connaissances et des croyances sur les comportements cognitifs humains, soit sur la manière d'apprendre et de traiter les informations apprises. Celles-ci peuvent être intraindividuelles, donc se référer à soi, interindividuelles, ou universelles (Flavell, 1987). Les connaissances cognitives liées aux personnes, réputées universelles, regroupent les connaissances de psychologie naïve (ou *folk psychology*), terme qui englobe un ensemble de croyances à propos du fonctionnement cognitif humain, par exemple que la mémoire a ses limites, ou encore que le système cognitif peut être en surcharge.

La seconde sous-catégorie concerne les connaissances des tâches. Dans le courant de leur développement cognitif, les individus en apprennent beaucoup sur la nature des informations et des contraintes qu'elles et ils doivent gérer. Par exemple, les tâches dans lesquelles les informations sont présentées de manière dense et non répétitive engendrent un traitement cognitif plus complexe. La réponse métacognitive à une tâche d'une telle nature sera plus lente et impliquera une activité de contrôle élevée. Les enfants qui savent prédire efficacement les demandes liées à une activité de compréhension de texte avec question de compréhension démontrent une connaissance accrue de la nature de la tâche (Flavell, 1985).

Le troisième sous-groupe de variables dans les connaissances cognitives concerne les stratégies ou les procédures qui permettent d'atteindre un but dans la réalisation

d'une tâche cognitive. Une stratégie métacognitive permet à l'individu de contrôler des stratégies cognitives dans l'atteinte de ce but (Flavell, 1985). Cette catégorie englobe aussi la connaissance des contextes qui demandent l'application de stratégies *cognitives* et de ceux qui requièrent des stratégies *métacognitives* (Giasson, 1990). Il peut s'agir par exemple d'additionner une liste de chiffres pour obtenir une somme, ou de faire une carte conceptuelle pour définir un concept. Lire plus ou moins rapidement en fonction de la qualité du texte traduit aussi la mise en place d'une stratégie cognitive. Par contraste, la lecture rapide, par balayage, d'un texte écrit, suivie d'une lecture approfondie, constitue une stratégie métacognitive de planification de la lecture, dans laquelle la tâche est d'abord évaluée avant d'être accomplie (Flavell, 1987).

Les *expériences* métacognitives sont des manifestations conscientes de l'utilisation de stratégies métacognitives (Flavell, 1987). Elles consistent en un contrôle et en la régulation de stratégies cognitives en fonction des informations mal comprises (Flavell, 1985). Elles sont ainsi marquées par exemple par le sentiment de ne pas comprendre quelque chose ou de faire une erreur. La prise de conscience et le désir de comprendre ces sentiments constituent la manifestation d'une expérience métacognitive, contrairement à la connaissance métacognitive, où il s'agit de juger de la difficulté inhérente à une tâche ou de l'utilisation de stratégies particulières pour pallier cet obstacle (Flavell, 1987).

L'expérience métacognitive en elle-même se caractérise par une pause, qui permet cette prise de conscience pertinente à la conduite d'une tâche ou de toute entreprise cognitive. Les jeunes enfants, devant une situation ambigüe, ou difficile à exécuter, remarquent l'ambiguïté de la tâche, mais sans comprendre la signification ou les implications de ces doutes, de ce bris de certitude[11] (Flavell, 1987). Dans ce cas, le jeune enfant ne sait pas comment rediriger son attention ni quelles stratégies mettre en place pour faire de cette expérience cognitive une expérience véritablement métacognitive. Ainsi, le doute préside à l'expérience métacognitive, qui donne alors lieu à la pause réflexive.

1.1.4 Synthèse du cadre sur la métacognition

Il ressort de notre revue succincte des travaux dans le champ de la métacognition deux concepts qui sont réinvestis dans la littérature métalinguistique, qui en est issue. Les habiletés métacognitives supposent la présence de ces éléments :

- Régulation (*monitoring*), soit la capacité de poser un jugement sur la qualité de ses processus cognitifs.
- Contrôle (*control*), soit l'utilisation de ces éléments d'évaluation pour produire des modifications de son comportement, par exemple l'inhibition des éléments non pertinents à la résolution de problèmes.

[11] De récentes découvertes empiriques soutiennent la thèse selon laquelle les représentations dans le *workspace* (sorte de mémoire de travail) comportent toujours une mesure de confiance, qui est assimilée à la métacognition, pouvant se traduire comme une évaluation d'un degré de certitude (Shea et Frith, 2019).

La pause, induite par le doute et par la prise de contrôle dans le processus de résolution de problème, constitue le lieu où les habiletés cognitives se manifestent.

1.2 Définitions de la conscience métalinguistique et statut des construits métalinguistiques

Les habiletés métalinguistiques, telles qu'elles sont définies au sein des travaux s'intéressant aux processus psychologiques et cognitifs à l'œuvre dans l'acquisition et l'utilisation des langues, sont comprises comme les habiletés à réfléchir (*monitor and control,* régulation et contrôle) sur les différentes unités constitutives du langage (Nagy, 2007), à voir le langage comme un code, en le séparant de sa signification symbolique (Friesen et Bialystok, 2012) ; « l'activité métalinguistique ne renvoie pas au langage sur le langage, mais à la cognition sur le langage et fait partie intégrante des activités métacognitives » (Gombert, 1990, p. 20). Partant, et comme mentionné précédemment, les HmL sont entendues comme un sous-domaine de la cognition et se réfèrent à des connaissances explicites sur différentes composantes de la langue, soit, entre autres, les phonèmes, les syllabes, les morphèmes et les groupes syntaxiques (Bialystok, 2001a; Gombert, 1990). Corollairement, on envisage la recherche sur les habiletés métalinguistiques selon l'élément linguistique en jeu : la syntaxe (habiletés métasyntaxiques), la phonologie (habiletés métaphonologiques), la morphologie (habiletés métamorphologiques), la sémantique (habiletés métasémantiques), le lexique (habiletés métalexicales) (Gombert, 1990).

La littérature sur les habiletés métalinguistiques a été investie par les chercheur·es dans les champs de l'acquisition du langage en langue seconde et première. Les recherches se classent généralement en deux grandes orientations[12]. La première concerne l'effet facilitateur du bilinguisme sur les HmL. Il s'agit d'une littérature surtout anglophone, marquée par les apports de Bialystok (1978, 1986, 1988b, 1993, 2001b)[13]. La deuxième grande orientation a comme objectif la prédiction des performances en lecture, ou d'autres compétences langagières, par les HmL en L1 et en L2 (Jeon, 2011; Simard, D. *et al.*, 2012a, 2012b). L'un des porte-étendards de ces travaux en langue première dans le monde francophone est Gombert (1988, 1990, 1996).

[12] La part de variance dans les différences interindividuelles va dans un sens ou dans l'autre selon l'orientation, ou la direction de l'hypothèse testée : dans l'une, le bilinguisme (VI) influe sur les HmL (VD), et dans l'autre, les HmL (VI) influent sur la compréhension en lecture (VD). Les conceptions théoriques influencent la nature des outils méthodologiques utilisés, ce qui peut contribuer à expliquer les confusions subséquentes dans l'interprétation de résultats.

[13] Les recherches sur l'effet facilitateur du bilinguisme sur le développement des habiletés métalinguistiques (Bialystok, 1988) se basent sur une idée proposée par Vygotsky dans *Pensée et langage* (1934/1997, cité dans Bialystok *et al.*, 2003; Malakoff et Hakuta, 1991). Un objectif sous-jacent à ce type de recherche est de vérifier si les bilingues sont à risque de moins bien développer certains aspects cognitifs ou langagiers par rapport aux monolingues. Le bilinguisme était réputé représenter une charge cognitive supérieure. La recherche a généralement montré un développement métalinguistique précoce chez les bilingues (Barac *et al.*, 2014), libellé « avantage bilingue ». L'hypothèse a posteriori était que le fait de connaître plus d'une langue induisait des performances supérieures du système exécutif, autant sur les éléments linguistiques que sur les éléments non linguistiques.
Les travaux de Bialystok (1988, 2001a, 2001b) ont permis de rendre compte et d'expliquer l' « avantage bilingue », mais surtout la variabilité dans les résultats de la recherche, par la diversité des méthodes, et plus précisément par la reconceptualisation des tâches utilisées pour mesurer les différents construits métalinguistiques.

Avant de procéder à une revue des modèles théoriques sur la conscience morphologique et la conscience syntaxique, il nous faut clarifier ce que nous entendons par « conscience métalinguistique » et « habiletés métalinguistiques », puisque ces concepts n'ont pas été utilisés de façon cohérente dans la littérature. L'adjonction du préfixe *méta-* à une pléthore de phénomènes provoque notamment des difficultés conceptuelles dont les ramifications s'étendent aux plans méthodologique et théorique (Bialystok, 2001a; Flavell, 1985; Gombert, 1996; Simard, D. *et al.*, 2017)[14]. Or, sans une forme de consensus terminologique, les variations interindividuelles sur le plan des habiletés métalinguistiques demeureront opaques (Bialystok, 2001b). C'est ce constat qui nous amène à revisiter les construits métalinguistiques, d'abord d'un point de vue théorique, en révisant les apports de la psychologie cognitive et de la didactique quant à cette question terminologique, et ensuite d'un point de vue méthodologique, en tenant compte du fait que les habiletés métalinguistiques s'incarnent dans les instruments de mesure mis en place pour les évaluer.

La section 1.2.1 est consacrée à la désambiguïsation des différents construits métalinguistiques et aux consensus terminologiques émergeant des modèles. Nous

[14] Flavell (1985) remet en question l'emploi du même suffixe pour définir les connaissances métacognitives, en ce que son utilisation suggère, à tort, que ce champ de connaissances est séparé, et plus « mystérieux » ou « spécial » que d'autres types de connaissances.
De la même façon, Bialystok (2001a, p. 121) lance une mise en garde sur la surutilisation du préfixe « méta » dans la littérature sur l'acquisition des langues et en éducation, en remarquant qu'il avait été utilisé, dans la littérature, pour définir de multiples « fonctions cognitives requérant des connaissances ou efforts augmentés » [notre traduction].

présenterons par la suite (1.2.2) les habiletés sous-jacentes pouvant être inférées des différentes définitions des HmL exposées, le modèle de Bialystok et Ryan (1985), développé et enrichi les années suivantes par Bialystok (2001a), étant le mieux à même de représenter ces dimensions.

1.2.1 Quelques consensus

La littérature abonde de définitions plus ou moins constantes de la conscience métalinguistique[15] (*metalinguistic awareness*). Il existe toutefois quelques éléments de convergence. Le premier consensus qui se dégage de notre revue de la littérature en psycholinguistique, en acquisition des langues et, plus généralement, en psychologie cognitive, concerne la différence fondamentale entre le « métalinguistique » et la « métalangue » (1.2.1.1). Le deuxième consensus porte sur le caractère « surplombant » des habiletés *méta*linguistiques, c'est-à-dire sur l'importance de cette prise de distance par rapport à l'objet linguistique dans la définition du « méta » (1.2.1.2). Le troisième consensus sur les caractéristiques définitoires des HmL passe par leur opposition à ce qu'elles ne sont pas : des habiletés linguistiques (1.2.1.3). Les HmL sont donc requises pour l'analyse et la

[15] C'est plutôt le terme « *capacité* métalinguistique » (« capacités de manipulation intentionnelle de la structure linguistique ») qui est utilisé par Gombert *et al.* (1994, p. 61). Les « exercices » métalinguistiques permettant à Gombert *et al.* (1994) de mesurer les capacités métalinguistiques recoupent effectivement les tâches typiquement utilisées dans la littérature sur la conscience phonologique et la conscience morphosyntaxique : dénombrement, suppression et permutation syllabiques, dénombrement, suppression et permutation phonémiques ; jugement, localisation, correction, réplication et explicitation d'erreurs ou agrammaticalités morphosyntaxiques (Gombert *et al.*, 1994).

manipulation des connaissances linguistiques spécialisées, mais pas dans le cadre d'une utilisation du langage quotidien, non spécialisée.

1.2.1.1 « Métalinguistique » n'est pas « métalangage »

Nous l'avons mentionné et démontré : la littérature scientifique sur les habiletés métalinguistiques, qu'elle soit didactique ou psycholinguistique, se heurte à des confusions importantes sur le plan terminologique en partie imputables à la productivité du préfixe *-méta*, qui se retrouve apposé à plusieurs noms représentant des concepts bien différents, et à la nature adjectivale du terme « métalinguistique » (Gombert, 1996) [16] . L'utilisation des couples de concepts « métalangage/métalinguistique » et « métalangue/métalangagier » est emblématique de ce phénomène (Boivin, 2009, 2014).

Il nous faut donc préciser que ce que nous entendons par « métalinguistique » appartient à la tradition psycholinguistique, et ne correspond pas au « métalangage », tel qu'il est utilisé dans la tradition linguistique et par extension, en didactique. Il s'agit d'une distinction importante, que relèvent notamment Benveniste (1974) et Gombert (1988, 1990, 1996). Les métalangues ou métalangages correspondent en effet, en linguistique, aux « langues qui servent à décrire une langue, dont c'est la seule et unique fonction […]. La langue de la grammaire, qui décrit l'usage des formes de la

[16] Le concept de « métalinguistique » est par ailleurs utilisé indistinctement pour qualifier les notions de *conscience* (awareness), d'*habileté* (ability) et de *tâche* (task) (voir Bialystok (2001a) pour une revue différenciée de ces construits).

langue, est une métalangue » (Benveniste, 1974, p. 35). Cette définition de la métalangue, en particulier dans son acception grammaticale, est abondamment reprise dans les travaux et les manuels didactiques (Reuter *et al.*, 2013), dans lesquels les connaissances grammaticales et leur explicitation sont reconnues comme des manifestations d'*activités* métalinguistiques (Boivin, 2009, 2014; Gauvin et Thibeault, 2016; Nadeau et Fisher, 2006; Simard, D. *et al.*, 2007). Le métalangage ne fait donc pas partie des caractéristiques définitoires des habiletés métalinguistiques dans le domaine de la psycholinguistique [17]. Gombert (1988, p. 65) rappelle l'importance de cette précision terminologique en contrastant le point de vue de linguiste à celui, béhavioral, de psychologue :

> […] le linguiste dépistera le métalinguistique en identifiant dans des productions verbales des marques linguistiques traduisant des processus d'autoréférenciation (utilisation du langage pour référer à lui-même), le psychologue quant à lui cherchera dans le comportement (verbal ou non) du sujet des éléments qui lui permettront d'inférer des processus cognitifs de gestion consciente, soit des objets langagiers en tant que tels, soit de leur utilisation.

Corollairement, et d'un point de vue plus pratique, il est aussi impératif de dissocier les *manifestations* d'habiletés métalinguistiques de l'utilisation du métalangage (Tunmer et Herriman, 1984) [18] : l'emploi de mots grammaticaux chez l'enfant n'est pas

[17] Les habiletés à verbaliser de façon formelle ou informelle des connaissances métalinguistiques, et donc à utiliser le métalangage grammatical, font toutefois effectivement partie d'instruments de mesure des habiletés métalinguistiques dans certaines études (voir à ce sujet les revues de Simard *et al.*, 2017, de Woll, 2019 et de Bourdages *et al.*, 2021).

[18] « Although metalinguistic awareness is related in meaning to the term *metalanguage,* it is important to distinguish between the two. While metalanguage refers to language used to describe language, and includes terms like *phoneme, word, phrase,* etc., metalinguistic

nécessairement le signe d'habiletés métasyntaxiques ou métagrammaticales élevées (Fortier *et al.*, 2019; Woll, 2019). De la même façon et inversement, une réflexion intentionnelle sur le langage peut être engagée sans qu'il y ait recours au métalangage ou à des règles grammaticales (Boivin, 2014; Gombert, 1990), et peut se manifester par l'application des manipulations syntaxiques ou la reproduction d'erreurs ciblées, par exemple dans Bourdages *et al.* (2021).

1.2.1.2 La langue comme objet

La *faculté* métalinguistique, telle que l'a définie Benveniste (1974, p. 228-229), se réfère à « la possibilité » que nous, comme êtres humains, sommes en mesure de déployer,« de nous élever au-dessus de la langue, de nous en abstraire, de la contempler tout en l'utilisant dans nos raisonnements et nos observations. [Cette faculté], à laquelle les logiciens ont été plus attentifs que les linguistes, est la preuve de la situation transcendante de l'esprit vis-à-vis de la langue dans sa capacité sémantique. »

Pour Malakoff et Hakuta (1991, p. 147), la conscience métalinguistique se définit comme « la conscience de la nature linguistique sous-jacente de l'utilisation de la langue. La conscience métalinguistique permet à l'individu de prendre du recul par

---

awareness refers to awareness of the *instantiations* of these terms, but not to knowledge of the terms themselves. Thus, a metalinguistically aware child may perform well on a task involving the manipulation of phonemes without knowing what the term *phoneme* means » (Tunmer et Harriman, 1984, p.12).

rapport à la compréhension ou à la production d'un énoncé afin d'examiner la forme et la structure linguistiques qui sous-tendent le sens d'un énoncé » [19] (notre traduction).

Gombert (1990) a suggéré que la conscience métalinguistique implique de réfléchir de façon intentionnelle au langage et de le manipuler. La langue elle-même devient un objet auquel on peut penser, qu'on peut objectiver. Le caractère délibéré de cette « conscience » est central dans sa définition. L'*activité* métalinguistique est un concept défini par Gombert (1990, 1996) comme une prise de conscience et une analyse des phénomènes langagiers, de manière à permettre leur contrôle et leur planification intentionnels en situation de compréhension et de production.

Sont ainsi comprises comme « métalinguistiques » les habiletés ou les activités se référant au langage et à son utilisation. Celles-ci se manifestent par des niveaux de conscience et d'explicitation plus élevés, par une appréhension surplombante des unités linguistiques. Il ressort de ces définitions que les objets linguistiques sont ainsi d'une part analysés, et d'autre part manipulés.

[19] «At the most general level, metalinguistic awareness may be defined as an awareness of the underlying linguistic nature of language use. Metalinguistic awareness allows the individual to step back from the comprehension or production of an utterance in order to consider the linguistic form and structure underlying the meaning of the utterance.»

1.2.1.3 Le « métalinguistique » par opposition au « linguistique »

Nous avons illustré à travers la recension de plusieurs définitions des construits métalinguistiques que celles-ci ont en commun cette prise de distance par rapport à un élément linguistique, qui devient objet d'analyse et de réflexion. Les habiletés métalinguistiques font ainsi référence aux habiletés permettant de réfléchir et de manipuler les caractéristiques structurelles du langage, en traitant le langage comme un objet de pensée. Elles rendent possible l'adoption d'une « attitude réflexive sur les objets langagiers et leur manipulation » (Gombert, 1990, p. 11). Cette habileté s'oppose, d'un point de vue définitionnel, à une utilisation du langage menant à la compréhension et à la production de phrases dans une langue donnée (Woll, 2019). Les habiletés métalinguistiques sont ainsi envisagées par le contraste avec les habiletés linguistiques, qui correspondent à un traitement langagier passif sur le plan de la production et de la compréhension.

Cette prise de position exige de poser la théorie sur le langage ou l'approche théorique sur la nature du langage à laquelle on souscrit. Celle-ci va en effet définir ce qui sera considéré comme « linguistique » et, corollairement, comme « métalinguistique ». L'option théorique générativiste, qui pose le langage comme compétence innée, suppose que tout travail sur la langue est *méta* un peu « par défaut » (Bronckart, communication personnelle, 2018), puisque ce travail métalinguistique, appris, reflète la mise en place d'habiletés qui varient sur le plan interindividuel et qui sont le produit d'une réflexion, ou, à tout le moins, d'un regard posé sur la langue qui n'est pas inné

(Bronckart, 2019). Cette approche théorique semble être celle qui domine dans les travaux sur les habiletés métalinguistiques, même si, à notre connaissance, les auteur·es des articles traitant de ce construit ne se positionnent pas formellement sur l'option théorique sous-jacente qui guide la conception de leur design expérimental et leur interprétation des données[20]. Ces travaux sont aussi en quelque sorte des artefacts d'une tradition psycholinguistique qui cherche à valider expérimentalement les conceptions chomskyennes par la reconnaissance de structures, indépendamment de leur sens (Le Ny, 2017).

Pour Tunmer et Herriman (1984),

> « […] la conscience métalinguistique peut être définie comme la capacité de réfléchir et de manipuler les caractéristiques structurelles du langage parlé, en traitant le langage lui-même comme un objet de pensée, par opposition à la simple utilisation du système linguistique pour comprendre et produire des phrases (Tunmer et Herriman, 1984, p. 12) [notre traduction] »[21].

Les auteurs prennent ici le pari, à l'instar de beaucoup d'autres (Bialystok, 2001a, 2001b; Gombert, 1996; Simard, D. *et al.*, 2017; Tunmer et Herriman, 1984), de définir

[20] Par contraste, l'enseignement de la grammaire de phrase dans les systèmes scolaires francophones est, depuis les années 1960-1970, pensé en accord avec la linguistique structuraliste et générativiste, et ce explicitement. Sont mis de l'avant le caractère systématique de la langue et la syntaxe et leurs régularités (Dolz et Simard, 2009).

[21] « *As a first approximation, metalinguistic awareness may be defined as the ability to reflect upon and manipulate the structural features of spoken language, treating language itself as an object of thought, as opposed to simply using the language system to comprehend and produce sentences.* »

le « métalinguistique » par opposition à ce qu'il n'est pas, soit le « linguistique » [22], en faisant ressortir le caractère conscient et actif de l'habileté métalinguistique, et le caractère inconscient et passif de l'activité linguistique.

1.2.1.4 La notion de conscience

Les habiletés métalinguistiques et leur caractère conscient se différencient des processus épilinguistiques (Culioli, 1968), soit ceux dont les sujets n'ont pas pleinement conscience (Gombert, 1990). Il s'agit d'un type de traitement qui se situe pour ainsi dire entre le *linguistique* et le *métalinguistique*. L'activité épilinguistique va au-delà de la réponse conditionnée, mais sans mobiliser des processus véritablement conscients et intentionnels (Culioli, 1968; Gombert, 1990, 1996). Culioli (1968, p.108), qui a été un des premiers à opérer cette distinction, définit de la façon suivante l'activité épilinguistique :

> « Le langage est une activité qui suppose, elle-même, une perpétuelle activité épilinguistique (définie comme « activité métalinguistique non consciente »), ainsi qu'une relation entre un modèle (la *compétence*, c'est-à-dire l'appropriation et la maitrise acquise d'un système de règles sur des unités) et sa réalisation (la *performance*) dont nous avons la trace phonique ou graphique, des *textes*. » (Les italiques sont de l'auteur.)

[22] Nous ne ferons pas référence au terme « épilinguistique » NI à ses concepts afférents dans le cadre de cette thèse, dans la mesure où ceux-ci nous semblent peu opératoires. En effet, on peut penser qu'une forme d'activité épilinguistique se manifeste lors de certaines activités linguistiques demandant des représentations intuitives sur la langue, mais ces représentations ne sont que difficilement mesurables. Les construits épilinguistiques ne sont par ailleurs pas explicitement évoqués dans le modèle du développement des habiletés métalinguistiques de Bialystok (2001a), qui constituera la pierre angulaire de notre réflexion.

Cette définition de l'activité épilinguistique comme une « activité métalinguistique non consciente » semble peu opératoire. Elle permet néanmoins à notre avis de tirer la chose suivante au clair : le « méta X » n'équivaut pas forcément à la « conscience de X ». Il ressort de cette inférence que les habiletés métalinguistiques ne sont pas nécessairement « conscientisées », ou « conscientes ». Elles peuvent être conscientes à un moment dans le développement langagier de l'individu, puis repasser sous le seuil de la conscience (au sens de Leontiev (1984)[23]). Par ailleurs, bien que relativement consensuelle sur le plan théorique, la manifestation de la conscience comme caractéristique définitoire des habiletés métalinguistiques est reconnue comme difficilement opérationnalisable (Fortier, 2013; Woll, 2019).

Gombert (1990) mobilise la notion de conscience dans son modèle des HmL (à laquelle il ajoute la dimension d'intentionnalité), alors que Bialystok (2001a) la rejette. Pour Bialystok (2001a), la conscience, en tant que processus *on-line*, est le produit de représentations mentales explicites et de l'attention. Les habiletés métalinguistiques (*metalinguistic awareness*) nécessitent que « l'attention soit activement dirigée vers les propriétés explicites de la langue »[24]. Il s'agit pour elle d'un phénomène dont la réalisation se déroule en un temps relativement court et relativement précis, puisque

[23] L'activité, pour Leontiev, permet d'opérer le passage d'éléments d'une phase de non-conscience à une phase consciente, par l'exercisation et par le travail. L'idée générale est de rendre la connaissance explicite au moment de la phase d'apprentissage, pour en arriver finalement à une phase où le transfert est automatisé.

[24] « *Metalinguistic awareness implies that attention is actively focused on the domain of knowledge that describes the explicit properties of language*. » (p. 126-127)

l'attention est dirigée sur des représentations mentales abstraites pour une raison précise, et mobilisée pour une résolution de problème momentanée (Bialystok, 2001a, p. 127). La conscience est donc indissociable de la théorie sur l'attention, et ce dernier construit permet précisément de porter le fardeau de la clarification terminologique beaucoup mieux que le concept « flou et controversé » qu'est la conscience (notre traduction de « *illusive and defiant* » dans Bialystok, 2001a, p. 128). L'auteure considère que l'attention est ce qui différencie le *métalinguistique* du *linguistique*, et préfère ce construit à la conscience. Outre l'attention, l'auteure invoque aussi l'opposition implicite-explicite pour caractériser les habiletés métalinguistiques.

Bialystok et Ryan (1985) soulignent que la conceptualisation traditionnelle de la conscience métalinguistique comme une habileté linguistique unique n'a pas été salutaire. Les auteures soutiennent que le terme « métalinguistique » devrait être appliqué « non pas à un accomplissement mental spécifique, mais plutôt à un ensemble de problèmes qui partagent certaines caractéristiques. La question théorique est donc de déterminer quelles compétences cognitives sous-tendent la solution à cet ensemble de problèmes » (Bialystok et Ryan, 1985, p. 230-231)[25]. » La

[25] « [*T]he term metalinguistic applies not to a specific mental accomplishment but rather to a set of problems which share certain features. The theoretical issue, then, is to determine what cognitive skills underlie the solutions to this set of problems* ». La citation se poursuit comme suit : «*Further, if other types of problems, such as literacy problems, can also be analyzed in this way, then commonality in the set of underlying skills necessary to deal with specific metalinguistic and specific literacy problems would indicate the areas of overlap in these problem type*», qui peut être traduite ainsi : «En outre, si d'autres types de problèmes, tels que les problèmes de litéracie, peuvent également être analysés de cette manière, alors les points communs de l'ensemble des habiletés sous-jacentes nécessaires pour traiter des problèmes métalinguistiques et de litéracie spécifiques indiqueraient les zones de chevauchement de ces types de problèmes » [notre traduction]. Cette dernière remarque est proche de l'impulsion qui

conscience métalinguistique – ou habiletés métalinguistiques – est ainsi traitée comme l'habileté qui permet d'aborder et de résoudre avec succès certains types de problèmes. Il s'agit ainsi à la fois d'une forme de conscience (*awareness*) et d'une habileté (*ability*), dans la mesure où le problème en lui-même est d'ordre métalinguistique, et l'habileté se situe dans la reconnaissance de la nature et des exigences liées au problème (Malakoff et Hakuta, 1991). Les notions d'habileté[26] et de conscience sont utilisées de manière presque interchangeable dans la littérature sur les habiletés métalinguistiques (Woll, 2019), ainsi que dans la présente thèse. Bialystok (1993, p. 217, cité dans Woll (2019, p. 50)) avance une clarification :

> *A learner performing a task classified as metalinguistic is demonstrating metalinguistic ability. [...] When applied to levels of awareness, virtually any performance can have metalinguistic properties if it is carried out with the deliberate control and awareness of the language learner.* (Bialystok, 1993, p. 217)

Cette clarification illustre pour nous le fait que la théorie métalinguistique et les outils méthodologiques métalinguistiques sont deux faces d'une même pièce, et renforce l'idée que le « méta » doit être évalué à l'aune des tâches qui sont censées le mesurer. Le modèle de Bialystok et Ryan (1985), que Bialystok (2001a) raffinera par la suite, donne corps à notre problématique, en fournissant un cadre théorique basé sur les

---

se trouve à l'origine de notre recherche et de sa démarche méthodologique : la découverte des habiletés sous-jacentes et communes entre les HmL et la compréhension en lecture.

[26] Le concept d' « habiletés » pose lui aussi un certain nombre de problèmes. Celui-ci doit être dissocié des « connaissances », qui renvoient à un type d'information en mémoire qui puisse être conscient ou non. Le construit « habiletés métalinguistiques » dénote l'accès à ces connaissances et la capacité à les utiliser (Simard, *et al.*, 2017).

dimensions constitutives de la conscience métalinguistique. La section suivante est dédiée à la description de ce modèle.

1.2.2 Dimensions sous-jacentes constitutives des habiletés métalinguistiques

Bialystok et Ryan (1985), puis Bialystok (2001a) ont proposé un modèle qui rend compte de la charge cognitive que représentent les tâches métalinguistiques les plus utilisées, afin de fournir un cadre d'analyse métacognitif au développement du langage en langue première et seconde.

Ce modèle (Bialystok, 1988b, 2001a, 2001b; Bialystok et Ryan, 1985) est dominant dans la littérature sur l'acquisition des langues secondes. Il met en avant une définition des habiletés métalinguistiques qui découle directement des outils qui servent à les mesurer, et dégage deux composantes procédurales qui sont réputées être mobilisées lors de la réalisation de tâches associées à ces habiletés; d'une part, *l'analyse des connaissances linguistiques*, représentée par l'axe des abscisses, et d'autre part, le *contrôle cognitif des traitements langagiers*, sur l'axe des ordonnées. Les éléments à gauche demandent un niveau d'analyse plus faible, et à droite, un niveau plus élevé ; les éléments en bas sont censés mobiliser des niveaux de contrôle plus faibles, et en haut, des niveaux plus élevés. Une mobilisation élevée des deux processus rend la tâche plus difficile, ce qui initie un comportement de plus en plus métalinguistique.

Cette conception à deux modalités, présentée à la figure 1.1, prévoit aussi qu'une tâche demandant de hauts niveaux de contrôle et d'analyse se situera dans le quadrant supérieur droit : plus la tâche mobilise ces composantes, plus elle est dite « métalinguistique » (Bialystok, 2001a; Simard, Fortier & Foucambert, 2011; Fortier, 2013). Le quadrant en bas à gauche permet de représenter les mesures qui mobilisent des habiletés linguistiques, ou de l'ordre du « conversationnel » (Bialystok, 1988b). Ainsi, ce modèle permet aussi de classer les tâches métalinguistiques et linguistiques selon les processus cognitifs exigés (haut ou bas niveaux), mais aussi, éventuellement, les activités grammaticales réalisées en salles de classe, en plus de rendre possibles des prédictions sur le développement langagier de l'élève. Le modèle princeps de Bialystok et Ryan (1985), présenté à la figure 1.1, présentait à l'origine les activités de lecture et d'écriture comme moins « métalinguistiques », ou comme nécessitant moins de contrôle attentionnel et d'analyse des connaissances que les habiletés métalinguistiques en elles-mêmes. La lecture et l'écriture sont aujourd'hui conçues comme des activités éminemment métalinguistiques.

Gombert (1990, p. 15) propose l'interprétation suivante de ces deux construits cognitifs (que nous expliciterons en 1.2.2.1 et 1.2.2.2) :

> L'analyse des connaissances est requise lorsque la situation de traitement linguistique est épurée des indices contextuels extralinguistiques qui permettent habituellement une production ou une compréhension non réfléchie. Le contrôle cognitif conscient intervient lorsque la situation demande une prise en compte privilégiée des aspects formels du langage aux dépens de la signification.

1.2.2.1 Analyse des connaissances linguistiques

L'axe qui représente l'analyse des connaissances linguistiques fait référence à la capacité à se représenter des structures de plus en plus explicites, de plus en plus fines et de plus en plus abstraites (Bialystok, 1986; 2001a, p. 131). Il s'agit du processus par lequel les représentations mentales implicites sont réorganisées en représentations explicites de la structure. Cette composante regroupe les habiletés de structuration, de réorganisation et d'explicitation des connaissances linguistiques (Cohen, 2011). Dans le cadre de cette conception du développement du langage, les connaissances implicites sont peu à peu transformées en connaissances explicites, qui sont conscientes et verbalisables (Karmiloff-Smith, 1986; Simard, D. *et al.*, 2017). Les connaissances explicites correspondent ainsi à la conséquence de l'analyse des connaissances linguistiques (Bialystok, 2001a; Bialystok et Ryan, 1985; Simard, D. *et al.*, 2017).

Des connaissances non analysées et implicites sont exigées régulièrement, au quotidien, notamment dans un contexte conversationnel. L'analyse des connaissances, quant à elle, est nécessaire dans les situations où la manipulation intentionnelle d'une règle, par exemple, est nécessaire. L'explicitation représente, dans ce cadre, un niveau maximal d'analyse des représentations linguistiques. Le développement de cette dimension apparait en même temps que la maturité cognitive générale : les connaissances implicites sont progressivement analysées en

catégories structurées [27] . Cet axe matérialise l'idée de prise de distance, d'objectivation de la langue, caractéristique des habiletés métalinguistiques (Fortier, 2013), tel qu'explicité en 1.2.1.2.

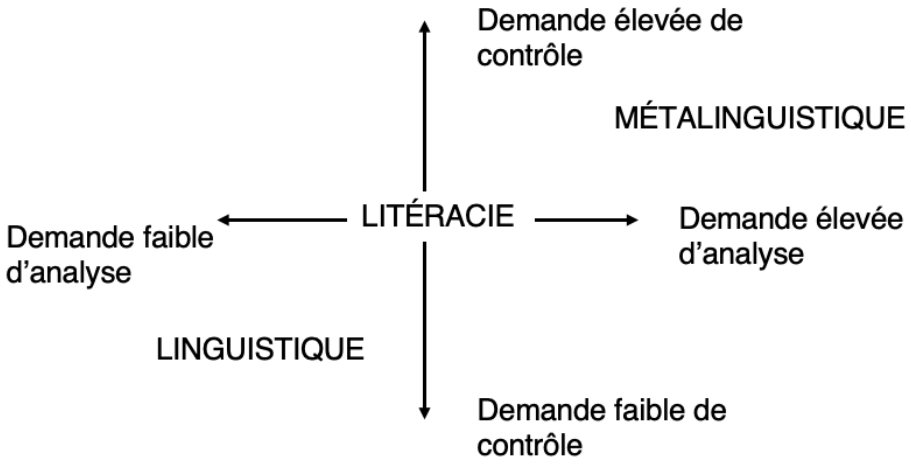


Figure 1.1 Le modèle de Bialystok et Ryan (1985, p. 232), puis de Bialystok (2001a, p. 16), tiré de la traduction de Fortier (2013, p. 22)

Bien que Bialystok (1986, 2001a) considère l'analyse des connaissances linguistiques comme une composante procédurale, nous sommes d'avis qu'elle constitue plutôt une

[27] L'opposition entre connaissances implicites et explicites est une notion assez largement partagée en psychologie cognitive (Huot et Schmidt, 1996 ; Bialystok, 1988b). L'usage des concepts implicite et explicite a essaimé dans la littérature, ce qui est source de confusion en raison de la variété des noms auxquels ils sont associés : « connaissance », « apprentissage », « mémoire », « grammaire », « condition » ou encore « règle » (Huot et Schmidt, 1996).
Les concepts de processus implicites et explicites sont par ailleurs définitoires des processus duaux (ou DPT, pour dual-process theory), théorie qui suppose l'existence de deux types de processus cognitifs mobilisés dans les tâches en psychologie cognitive et en psychologie sociale : des processus rapides, automatiques et implicites (type 1), et des processus analytiques, délibérés, explicites (type 2) (Thompson *et al.* 2011 ; Evans et Stanovich, 2013). Nous mettrons ces processus en lien avec les habiletés métalinguistiques dans le chapitre X.

habileté de type déclaratif, surtout dans la mesure où sa représentation maximale correspond à une explicitation formelle des représentations mentales liées à ces connaissances (Karmiloff-Smith, 1986; Simard, D. *et al.*, 2017). Il nous semble, par ailleurs, que c'est à cette condition que l'analyse et le contrôle peuvent être représentés comme deux dimensions orthogonales.

1.2.2.2 Contrôle cognitif

Le contrôle[28] des traitements langagiers (Tunmer et Herriman, 1984) a été défini par Bialystok (2001a) comme la composante des habiletés métalinguistiques qui permet de diriger l'attention intentionnellement sur la sélection et l'intégration des informations pertinentes et appropriées (Bialystok, 1988b, p. 561) et la capacité de s'occuper sélectivement d'aspects spécifiques d'une représentation, particulièrement dans des situations trompeuses ou ambigües, et qui nécessitent une solution (Bialystok, 2001a, p. 131). Cet axe matérialise, quant à lui, la notion d'attention volontaire et délibérée, définitoire des HmL chez Gombert (1990), notamment.

[28] La notion de contrôle de l'attention se retrouve aussi dans *Pensée et langage* (Vygotski, 1934/1997). Vygotski considère cette habileté à contrôler l'attention de façon délibérée comme l'une des fonctions à la base de la formation des concepts dans le développement de l'individu. « Le concept […] apparait dans le processus d'une opération intellectuelle ; ce n'est pas le jeu des associations qui entraine la construction d'un concept : à sa formation participent toutes les fonctions intellectuelles élémentaires en une combinaison spécifique, l'élément central de cette opération étant l'emploi fonctionnel du mot comme moyen de diriger volontairement l'attention, d'abstraire, de différencier des traits, d'en faire la synthèse et de les symboliser à l'aide d'un signe » (Vygotski, 1934/1997, p. 267).

Situé dans le domaine procédural (Fortier, 2013), le contrôle cognitif représente une fonction exécutive proche de la notion de mémoire de travail (Bialystok et Ryan, 1985), responsable de la sélection et de la coordination de l'information lors de son traitement. Le décours temporel associé à la sélection d'une information pertinente est un indice de contrôle plus ou moins bien maitrisé. Si de nombreuses sources doivent être consultées, la coordination est difficile. Ce construit cognitif, qui renvoie à la notion de contrôle conscient (*conscious control*), s'applique aux locuteurs natifs et non natifs de langues (Simard, D. *et al.*, 2017).

Certaines tâches métalinguistiques mobilisant des capacités de contrôle plus élevées peuvent exiger, par exemple, que les sujets concentrent leur attention sur certains éléments linguistiques, tout en ignorant ou en inhibant une compréhension du sens ou d'informations entrant en contradiction avec le but de la tâche et le construit visé (Cohen, 2011).

1.2.2.3 Représentations et processus

Il nous semble important ici de situer les habiletés métalinguistiques sur le plan épistémologique. L'habileté métalinguistique, pour une unité linguistique X (par exemple, un morphème) implique que X est à la fois l'objet d'une attention intentionnelle et la cause de la prise de conscience (Tunmer et Herriman, 1984, p. 12). Dans cette conception, et si l'on admet de concevoir les HmL et ses habiletés sous-jacentes avec un niveau de granularité assez élevé, l'unité linguistique X existe

comme partie intégrante du système linguistique, ou comme représentation linguistique, et peut devenir l'objet d'une réflexion abstraite, consciente, par un processus que nous avons appelé contrôle cognitif.

Nous revenons ici vers deux notions fondamentales et prédominantes dans la littérature psycholinguistique : celles de *représentation* et de *processus* (Le Ny, 2017), et nous situons ainsi dans un cadre théorique qui conçoit la faculté de langage comme non modulaire, ou reposant sur des mécanismes cognitifs généraux (*domain-general*). Les représentations correspondent à des unités de contenu (existant en mémoire de travail ou en mémoire à long terme), et les processus relèvent du traitement de l'information, ou de la modification que subissent les représentations (Le Ny, 2017). La description des différentes représentations mentales permet de faire la lumière sur les bases cognitives du langage, et l'analyse des processus mentaux permet de répondre à la question du traitement du langage au sein du système cognitif (Le Ny, 2017).

« En termes de traitement de l'information, les opérations métalinguistiques peuvent être décrites comme des opérations mentales effectuées sur les sorties des sous-systèmes modulaires impliqués dans la compréhension des phrases » (notre traduction de Tunmer (1997, p. 28))[29]. Ces sous-systèmes incluent la phonologie, la morphologie et la syntaxe.

[29] Notre traduction de « *In information processing terms, metalinguistic operations can be described as mental operations performed on the outputs of the modular subsystems involved in sentence comprehension* » (Tunmer, 1997, p. 28).

Les HmL se déclinent donc sur des unités de traitement – morphologique (Carlisle, 2000; Casalis *et al.*, 2009) ou syntaxique (Foucambert, 2008; Gaux et Demont, 1997), notamment. Mais la réussite aux épreuves mesurant les différentes HmL requiert également une habileté dans l'opération spécifique qui doit être réalisée, indépendamment de son contenu. Ce constat nous amène à nous représenter alors deux grandes familles d'habiletés latentes et orthogonales de nature bien différente : l'une est de nature linguistique, tandis que l'autre est de nature cognitive.

1.3 Synthèse du cadre sur les habiletés métalinguistiques

Cette revue de la littérature sur les habiletés métalinguistiques nous permet d'insister finalement sur deux éléments constitutifs propres à la définition du *métalinguistique* (Demont *et al.*, 2006; Fortier, 2013, p. 13; Gombert, 1990) qui seront réinvestis dans la typologie des épreuves qui les mesurent : d'abord, la *langue comme objet de l'attention volontaire*, ensuite, le *contrôle et la manipulation* des éléments langagiers. Ce sont ces caractéristiques qui nous permettront de mieux comprendre les dimensions sous-jacentes mises en œuvre dans les tâches, et ultimement, comment et pour quelles raisons la conscience morphologique et la conscience syntaxique, et par extension, la conscience phonologique, permettent de prédire les performances en lecture.

Par conséquent, la conscience morphologique sera définie comme les habiletés à reconnaitre, à manipuler et à contrôler délibérément les structures morphologiques, et la conscience syntaxique sera définie comme les habiletés permettant de reconnaitre, de manipuler et de contrôler intentionnellement les structures syntaxiques (Nagy, 2007). Ainsi, la conscience morphologique et la conscience syntaxique – parfois réunies sous le vocable *conscience grammaticale* (Bowey, 1994; Tong *et al.*, 2014) – elles-mêmes des sous-branches des HmL –, consistent en l'identification et la manipulation consciente des unités linguistiques auxquelles elles se rapportent (Gombert, 1990). Au sein du système de traitement du langage, les habiletés en conscience syntaxique interagissent certainement avec celles se rapportant à la conscience morphologique (MacKay *et al.*, 2021). Ces deux faces de la conscience grammaticale sont réputées régir notamment les règles d'accord et celles du choix des temps verbaux, par exemple, se situant à l'interface entre morphologie et syntaxe.

Nous avons par ailleurs situé les HmL comme habiletés métacognitives. La conscience morphologique et de la conscience syntaxique, autrement appelées habiletés métasyntaxiques et habiletés métamorphologiques, s'étudient dans ce cadre plus général des habiletés métalinguistiques, ou conscience métalinguistique. La revue de la littérature nous a permis d'asseoir l'intérêt de présenter les construits psycholinguistiques en les décomposant en fonction des représentations et des processus qu'ils mettent en œuvre (Le Ny, 2017), ce qui guide nos choix

méthodologiques présentés au chapitre VI et l'analyse de nos données au chapitre VII.

Dans les chapitres II et III, qui portent sur la conscience morphologique et la conscience syntaxique, nous appréhenderons les deux construits à travers le prisme des outils méthodologiques servant à leur mesure, dans le prolongement des efforts de Bialystok et Ryan (1985). Nous procéderons à un examen de ces tâches, considérées comme métalinguistiques, en fonction des prérequis cognitifs et linguistiques indispensables à leur résolution.[30] Notre objectif sera de développer une compréhension plus fine des enjeux liés au choix des outils de mesures et des habiletés sous-jacentes qu'elles mettent en œuvre afin de faire la lumière sur la relation entre les habiletés métalinguistiques et la lecture.

[30] « *To clarify what is intended by the term metalinguistic ability, a variety of tasks, considered to be metalinguistic, are examined in terms of the cognitive and linguistic prerequisites to their solution »* (Bialystok et Ryan, 1985, p.229).

Il était grilheure ; les slictueux toves
Gyraient sur l'alloinde et vriblaient :
Tout flivoreux allaient les borogoves ;
Les verchons fourgus bourniflaient.

— Lewis Carroll, *Alice au pays des merveilles*

CHAPITRE II

LA CONSCIENCE MORPHOLOGIQUE

Nous avons, dans le chapitre I, effectué une revue de la littérature sur les habiletés métalinguistiques et le cadre de recherche général dans lequel elles s'intègrent : celui des habiletés métacognitives. Cette revue nous a amenée à désambiguïser certains construits du champ et à asseoir le cadre terminologique et théorique dans lequel cette recherche s'inscrit. Le présent chapitre sera consacré à l'étude d'un des deux grands construits métalinguistiques nous intéressant plus particulièrement : la conscience morphologique. Après une présentation du construit et de son développement (2.1), nous passerons en revue les épreuves étant utilisées pour mesurer la conscience morphologique et les dimensions sous-jacentes qu'elles mettent en œuvre dans la partie 2.2. Nous ferons un état des lieux des liens directs et indirects entre la conscience morphologique et la compréhension en lecture (2.3) et clôturerons le chapitre par une synthèse (2.4).

2.1 Représentations morphologiques

La morphologie constitue un domaine de la linguistique qui étudie la façon dont se combinent les morphèmes, entendus comme les unités minimales de sens, dans une langue donnée (Colé *et al.*, 2004; Kuo et Anderson, 2006). Les unités morphologiques permettent d'établir les relations sémantiques ou syntaxiques entre les mots et les propositions (Carlisle, 1995; Chomsky et Halle, 1968). Les morphèmes sont dits libres lorsqu'ils agissent comme mots autonomes, comme *avenue* ou *beau.* Ceux-ci peuvent alors fonctionner comme bases. On retrouve aussi la catégorie des morphèmes liés (Anderson, 2018). Ceux-ci ne peuvent fonctionner comme base ou comme mot indépendants : ils doivent être adjoints à des bases, c'est-à-dire qu'ils fonctionnent par affixation, comme dans « beau*té* » et « avenue*s* ».

Les morphèmes flexionnels et dérivationnels constituent les deux sous-types de morphèmes liés du français. La morphologie dérivationnelle renvoie aux mots composés d'au moins une base et d'un ou plusieurs affixes qui expriment une information sémantique et qui peuvent modifier la catégorie grammaticale et/ou la signification sémantique de la base (Bloomfield, 1933; Colé *et al.*, 2004). Le mot plurimorphémique « lion*ceau* » est composé du suffixe *-eau* (plus précisément de son allomorphe[31] *-ceau*) qui, adjoint à la base *lion*, permet de composer le sens de 'petit du lion' ou encore de 'petit lion', en fonction du contexte. De façon relativement

[31] Un allomorphe est une variante phonétique (qui se traduit le plus souvent par une altération orthographique) d'un même morphème qui n'en altère pas le sens.

analogue, *re*lire est formé de la base *lire* et du préfixe *-re*, ce dernier dénotant une idée de répétition, pour former un sens global de 'lire à nouveau'. Les morphèmes flexionnels fournissent quant à eux des informations grammaticales additionnelles sans modifier la catégorie de la base. Les mots fléchis sont ainsi constitués d'une base et d'une ou plusieurs flexions, marquant, à l'écrit, le genre et le nombre (seul, seul*e*, seul*es*), et dans le cas des verbes en français, le temps et la personne (ex. oubli*ez*, oubli*erez*) (Colé *et al.*, 2004; Muse, 2005)[32].

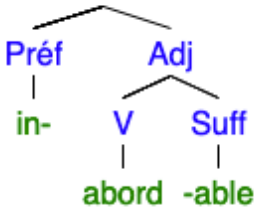
En morphologie générative, les morphèmes individuels des mots multimorphémiques, ou complexes, sont traditionnellement représentés de manière hiérarchique avec branchement binaire (Song *et al.*, 2019). L'analyse du mot *inabordable*, par exemple, permet d'exprimer le caractère compositionnel de la morphologie : ce mot multimorphémique avec branchement à droite est formé à partir de la base verbale *abord(er)*, à laquelle est ajouté le suffixe dérivationnel *–able* ('caractère de ce qui est X'), formant la nouvelle base *abordable*, à laquelle se greffe le préfixe *in–* ('n'est pas X'). Cette analyse de « inabordable », signifiant 'qu'il n'est pas possible d'aborder' est représentée en (1), (2) et (3).

(1) [in-[[abord]$_{V}$-able]] $_{ADJ}$] $_{ADJ}$

(2) [[Préf[in-]][Adj[V[abord]][Suff[-able]]]

[32] Les flexions casuelles sont marquées dans certaines langues, notamment en allemand.

(3)



Cette représentation arborescente domine largement dans la littérature morphologique, notamment en raison de son rôle supposé dans le traitement de la langue, en particulier dans le calcul de la signification lexicale (Lavric *et al.*, 2011). Ainsi, dans l'analyse de mots plurimorphémiques, le tout est la somme de ses parties, et l'accès au sens se fait par l'entremise d'une analyse plus ou moins réussie et plus ou moins automatique de chacune des composantes du mot (Taft, 1981)[33]. Pendant le traitement de mots morphologiquement complexes, les représentations du mot entier ainsi que ses constituants (base et affixes) sont activés (El-Bialy *et al.*, 2013). Les données empiriques montrent que les performances dans les tâches de reconnaissance lexicale, ou encore de dérivation, par exemple, sont significativement moins élevées plus le lien entre une base et son dérivé est opaque (Carlisle et Nomanbhoy, 1993). La transparence d'un mot multimorphémique se caractérise par

[33] Les données de travaux sur la reconnaissance lexicale de mots fréquents notamment, indiquent que les mots vont être perçus dans leur globalité (Butterworth, 1983) en fonction notamment de leur fréquence, suggérant ainsi que la forme d'un mot multimorphémique peut être perçue d'abord, plutôt que ses parties constitutives (Ferrand et Spinelli, 2005). Cet état de choses fait écho à certains principes de la gestalt et la perception, qui reposent sur des regroupements perceptuels des inputs engendrés par l'environnement (Rock et Palmer, 1990).

une transformation phonologique nulle de la base en cas d'affixation (exemple : *pomme – pommier*, au contraire de *pluie – pluvieux*).

2.1.1 Conscience morphologique

Les connaissances des représentations morphologiques sont susceptibles d'être mobilisées lorsque les habiletés en conscience morphologique sont mises en œuvre (Bratlie *et al.*, 2022). La conscience morphologique se définit traditionnellement comme l'identification et la manipulation conscientes de la structure morphologique des mots (Carlisle, 1995; Gombert, 1990). Cette habileté à repérer avec succès la structure morphologique des mots permet, par ailleurs, de discriminer les mots en fonction de leur nature (Roy et Labelle, 2007) – adjectifs, noms, verbes, adverbes – et contribue aux performances en lecture de mots isolés (Carlisle, 1995, 2000; Muse, 2005).

La conscience morphologique implique de prendre de la hauteur par rapport aux unités morphologiques, de les objectiver. Concrètement, cette habileté se déploie en compréhension et en production, dans des situations où il s'agit de comprendre les changements significatifs que l'apposition d'un *-s* final à un nom ou d'un *anti-* à un adjectif fournit comme information supplémentaire. Il peut s'agir aussi de la conscience des changements orthographiques induits sur certains affixes, lorsqu'apposés à une base, comme le dédoublement ou la suppression de consonnes (Apel *et al.*, 2021).

2.1.2 Développement de la conscience morphologique

Les enfants commencent à utiliser des morphèmes liés de manière spontanée en production peu de temps après le début de l'utilisation de mots ou morphèmes libres, vers deux ans (Shi, 2014). L'acquisition des morphèmes flexionnels et dérivationnels est marquée par des différences sur le plan développemental. L'acquisition des marques de flexion a été massivement évaluée en anglais, grâce au protocole du *wug test* (Berko, 1958). Les flexions sont réputées être bien maitrisées assez tôt, souvent dans les premières années de l'école primaire (Apel *et al.*, 2013; Berko, 1958; Casalis et Louis-Alexandre, 2000; Roy et Labelle, 2007), soit avant même l'apprentissage formel de la lecture. La morphologie flexionnelle, qui englobe les marques casuelles, de nombre, de genre, de temps et de personne, correspond à une classe fermée et limitée de suffixes grammaticaux très fréquents. La morphologie dérivationnelle, quant à elle, qui comprend un assez grand nombre de préfixes, de suffixes et de mots composés, est une classe ouverte (Apothéloz, 2002; Casalis et Louis-Alexandre, 2000). Elle continue à se développer après le passage par l'école primaire, chez certains élèves (Roy et Labelle, 2007).

Les travaux sur le développement de la conscience morphologique se font assez rares, de sorte que les informations que nous avons à ce sujet sont parcellaires (Berthiaume *et al.*, 2010)[34]. Les enfants d'âge préscolaire peuvent manifester une

[34] La présente revue de la littérature sur le développement métamorphologique inclut un nombre important de travaux comportant des sujets anglophones, et est la conséquence de l'abondance de recherches en langue anglaise sur ce thème. Nous tenons ainsi pour acquis

forme de conscience morphologique, bien qu'elle soit quelque peu implicite (Colé *et al.*, 2004). Le développement des connaissances morphologiques apparait par ailleurs avec la maturité cognitive générale : les connaissances intuitives sont progressivement conscientisées et analysées en catégories structurées et explicites (Casalis et Louis-Alexandre, 2000; Duncan *et al.*, 2009). Entre 5-6 ans et 7-8 ans, les sujets démontrent une conscience progressivement plus forte de la relation entre les bases et leurs formes fléchies et dérivées (comme la relation entre *slow* et *slowest*) (Apel *et al.*, 2013).

Les connaissances morphologiques dérivationnelles se scindent en quatre aspects : relationnel, syntaxique, distributionnel (Kuo et Anderson, 2006; Tyler et Nagy, 1989) et réceptif (Roy et Labelle, 2007). Les connaissances sur le plan relationnel consistent en une reconnaissance de la structure interne complexe des mots et du fait que ceux-ci puissent partager une base commune. Les tests de jugement de relations de mots (comme l'épreuve de paires de mots, décrite au chapitre VI) sont un bon exemple

---

que le développement des sujets francophones ne diffère pas fondamentalement de leurs pairs anglophones en ce qui a trait à la conscience morphologique et la conscience syntaxique. Toutefois, mentionnons que Duncan *et al.* (2009) ont démontré que le développement de la conscience de la morphologie dérivationnelle est accéléré chez les francophones, qui déploient par ailleurs une plus grande habileté à généraliser les connaissances morphologiques à de nouveaux contextes par rapport aux anglophones, un phénomène que les auteures appellent le *French advantage*. Cet état de choses, pour elles, peut être expliqué d'abord par la productivité du système de formation des mots du français sur les plans flexionnel et dérivationnel (De Clercq et Housen, 2019). L'effet de l'impact des différences fondamentales entre les systèmes morphologiques des langues sur le développement de la conscience morphologique est corroboré notamment par Zhang *et al.* (2011) chez des sujets sinophones et anglophones. Le *French advantage* pourrait aussi être un artefact des différences entre les pays francophones sur le plan éducatif. En effet, il semblerait que le programme scolaire français introduise les notions de morphologie plus rapidement que dans le programme anglais (à ce sujet, voir aussi Redmond *et al.,* 2021).

d'épreuves mesurant cet aspect (Muse, 2005; Roy et Labelle, 2007). La connaissance syntaxique constitue l'habileté à identifier que les suffixes dérivés déterminent une catégorie syntaxique. Il s'agit ainsi de connaitre, par exemple, que le suffixe *-ité* est apposé à des bases souvent adjectivales pour former des noms, comme dans *neutralité*. Cet aspect est mesuré notamment par les tâches de sélection de mots dérivés, comme la tâche de choix de pseudomots dérivés (cf. section 6.2.1.5) (Kuo et Anderson, 2006). L'aspect distributionnel regroupe les connaissances des contraintes à l'œuvre dans l'adjonction des suffixes aux bases, c'est-à-dire que *-ité*, par exemple, peut s'adjoindre à des bases adjectivales, mais pas à des bases verbales ou nominales (Muse, 2005; Roy et Labelle, 2007; Tyler et Nagy, 1989). La tâche d'extraction de la base (6.2.1.2) et les épreuves connexes de type « plausibilité lexicale » sont susceptibles de mesurer ce type de connaissances. Les connaissances morphologiques réceptives concernent les affixes et leur fréquence. Il s'agit d'une connaissance intuitive, qui permet de déterminer que *-able* est un suffixe fréquent, alors que *-upe* l'est beaucoup moins (Roy et Labelle, 2007).

Le développement de ces quatre aspects n'est pas simultané (Muse, 2005). Cependant, les connaissances syntaxiques et distributionnelles présupposent des connaissances relationnelles établies. Tyler et Nagy (1989) ont ainsi proposé l'ordre d'acquisition suivant, qui a été validé chez des élèves francophones par Roy et Labelle (2007) : connaissance réceptive – relationnelle (développée avant la 4e année, vers 9-10 ans) – syntaxique – distributionnelle (pas encore pleinement développée vers la

8e année, vers 13-14 ans). Les sujets de la recherche de Roy et Labelle (2007) avaient quant à eux entre 6 ans 8 mois et 8 ans 7 mois[35].

Les enfants semblent acquérir une connaissance rudimentaire de la morphologie dérivationnelle relationnelle, c'est-à-dire la capacité de reconnaitre une base familière et ses dérivés, avant la quatrième année d'études (soit vers 9-10 ans). La connaissance des propriétés syntaxiques des suffixes dérivés semble augmenter jusqu'à la huitième année d'études (vers 13-14 ans) (Tyler et Nagy, 1989). La connaissance des propriétés distributionnelles des suffixes augmente également, les élèves de sixième année présentant une augmentation des erreurs de surgénéralisation parallèle à celle constatée pour les suffixes flexionnels chez les plus jeunes (Muse, 2005).

Les grandes tendances développementales qui se dégagent des recherches citées peuvent constituer des artefacts des épreuves qui servent à mesurer les différentes connaissances relevées, et de la difficulté associée. Les tâches de type « jugement de relation de mots » sont souvent décrites comme faciles, puisqu'elles font appel à des connaissances implicites (Carlisle, 1995). Pourtant, elles pourraient être rendues complexes par le recours à des items sémantiquement opaques, comme *toile – toilette*, ou phonologiquement opaques, comme *fou – folie*. Ainsi, la conscience morphologique est susceptible de se développer tout au long de la vie, comme en

[35] 1re et 2e année primaire, au Québec.

témoignent notamment les recherches menées par Tighe, notamment, sur la conscience morphologique chez les adultes avec un faible taux de litéracie (Tighe, 2015; Tighe et Binder, 2015; Tighe *et al.*, 2018; Tighe et Schatschneider, 2015; Tighe et Schatschneider, 2016), celle de Wilson-Fowler et Apel (2015) chez des étudiant·es de niveau collégial, ou encore celle de Guo *et al.* (2011) chez des adultes.

2.2 La mesure de la conscience morphologique

En prenant comme base la définition communément admise de la conscience morphologique, soit l'habileté à repérer et à manipuler des morphèmes (Carlisle, 1995), les tâches utilisées dans les recherches sur la compréhension évaluent l'habileté des sujets à ajouter ou à supprimer un morphème d'un mot ou de pseudomots, avec contexte phrastique, ou sans (Tong *et al.*, 2014).

La littérature méthodologique portant sur la conscience morphologique est marquée par d'importantes disparités. La plupart des analyses ayant été effectuées sur la mesure de ce construit théorique signalent en effet unanimement la diversité des tâches qui sont utilisées pour le circonscrire (Apel, 2014; Apel *et al.*, 2013; Casalis *et al.*, 2002; Deacon *et al.*, 2006; Deacon *et al.*, 2008)[36].

[36] D'après nos recherches, un seul instrument de mesure, dont la qualité psychométrique et la fiabilité sont éprouvées, est disponible, en anglais (cf. *Morphological Awareness Test for Reading and Spelling* (MATRS)) (Apel, *et al.*, 2021).

2.2.1 Dimensionnalité de la conscience morphologique

La grande variété des tâches permettant de mesurer la conscience morphologique a poussé les chercheur·es à les reconceptualiser. Certain·es l'ont fait d'un point de vue théorique (cf. Berthiaume *et al.* (2010); Deacon *et al.* (2008)) rappelant la démarche méthodologique princeps de Bialystok (2001a) sur les processus impliqués dans les performances différenciées des personnes monolingues et bilingues dans les tâches de conscience linguistique (analyse et contrôle : voir la section 1.2.2 pour une revue de la littérature à ce sujet). D'autres l'ont testé empiriquement – voir notamment Bourdages et Foucambert (2018); Goodwin *et al.* (2017); Muse (2005); Spencer, Mercedes *et al.* (2015); Tibi et Kirby (2017); Tighe et Schatschneider (2015) –, en déterminant expérimentalement la validité du construit (*construct validity*) des épreuves de conscience morphologique, notamment par le biais d'analyses factorielles et confirmatoires. Au coeur de cette démarche se trouve la formulation d'hypothèses sur la dimensionnalité sous-jacente des mesures de la conscience morphologique, donc des facteurs de variation au sein même des tâches.

La dimensionnalité de la conscience morphologique a fait l'objet de travaux en anglais (Goodwin *et al.*, 2017; Muse, 2005; Spencer, Mercedes *et al.*, 2015; Tighe et Schatschneider, 2015), en arabe (Tibi et Kirby, 2017) et en français (Bourdages et Foucambert, 2018). Plusieurs études ont confirmé que la CM était mieux représentée comme un construit unidimensionnel (Muse, 2005; Spencer, Mercedes *et al.*, 2015). L'hypothèse de l'unidimensionnalité de la conscience morphologique suggère qu'il

s'agit d'un construit cognitif uniforme qui s'observe et se mesure sans égard à la nature des épreuves utilisées. A contrario, les représentations bidimensionnelles ou multidimensionnelles postulent que la conscience morphologique est composée de plusieurs facteurs ou sous-parties représentant des habiletés cognitives et linguistiques variables, et que sa mesure se montre sensible à cette variation. Tighe et Schatscheider (2015) ont effectivement montré que la conscience morphologique était mieux représentée comme un construit bidimensionnel impliquant de vrais mots et des pseudomots, tandis que dans l'étude de Tibi et Kirby (2017), le modèle bidimensionnel impliquant la modalité de la tâche (oral vs écrit) permettait de mieux circonscrire le phénomène. Goodwin *et al.* (2017) ont montré qu'un modèle bifactoriel comportant un facteur général représentant les connaissances morphologiques générales et sept facteurs spécifiques distincts représentant des tâches qui exploitent différentes facettes de la connaissance morphologique, qui ne sont pas subsumées par cette connaissance générale, permettait de mieux représenter la variation au sein des résultats aux épreuves de conscience morphologique.

Nous ferons état de ces facteurs de variation par une taxonomie des tâches utilisées pour mesurer la conscience morphologique dans la littérature[37].

[37] Cette taxonomie est notamment parue dans Bourdages et Foucambert (2018). Certaines parties du texte ont été reprises intégralement ici.

2.2.2 Taxonomie des épreuves de conscience morphologique

L'évaluation de la conscience morphologique est effectuée à l'aide d'épreuves variées et contrastées, et par conséquent, les habiletés qu'elles sont censées mobiliser diffèrent aussi (Casalis *et al.*, 2002; Goodwin *et al.*, 2017; Muse, 2005). Il nous faut considérer les facteurs de variation dans les épreuves qui sont réputées mesurer le même construit théorique – c'est-à-dire la conscience morphologique, dans le cas qui nous concerne. On retrouve notamment dans la littérature des épreuves de jugement de relation de mots, de jugement de grammaticalité, d'identification de l'erreur suivie d'une correction, de production de dérivés pour une même base, d'identification d'affixes dans des pseudomots (ex. *mouteur*) ou encore de complétion de phrases à l'aide de pseudomots (Kieffer *et al.*, 2016; Roy et Labelle, 2007).

2.2.2.1 Modalités de présentation

De prime abord, les tâches se différencient sur le plan de la modalité de leur présentation à l'élève (*input*), qui peut fort bien être différente de la modalité de la réponse demandée (*output*) (Deacon *et al.*, 2008). Les habiletés en décodage, notamment, sont susceptibles d'influencer les résultats dans les temps de réponse, et même dans la qualité de la réponse fournie. L'orthographe a l'avantage de maintenir les indices morphologiques au sein des items, ce qui pourrait faciliter la passation des épreuves pour certain·es participant·es. Il est important de rappeler, comme le soulignent Deacon *et al.* (2008), que dans le cas où les tâches sont présentées et recueillies à l'écrit, les morphèmes constituent aussi des unités orthographiques : il

s'agit donc d'un facteur qui doit être contrôlé ou analysé.

2.2.2.2 Types de morphèmes (unités)

Le traitement des dérivations est réputé être plus complexe que celui des flexions, car le procédé de dérivation implique souvent une altération phonologique et orthographique de la base, en plus de rendre possible une transformation de la catégorie du mot (Casalis *et al.*, 2002; Tighe et Schatschneider, 2015). L'acquisition des connaissances sur les morphèmes flexionnels et dérivationnels est différée, d'un point de vue développemental; comme mentionné en 2.1.2, les flexions sont maitrisées avant la morphologie dérivationnelle, qui continue à se développer au-delà des premières années d'école (Roy et Labelle, 2007). Plusieurs tâches de conscience morphologique n'impliquent que des morphèmes flexionnels ou des morphèmes dérivationnels; d'autres impliquent une combinaison des deux. Cependant, il reste à évaluer de façon empirique si ces deux types d'unités représentent des facettes distinctes de la conscience morphologique.

2.2.2.3 Opérations

Apel (2014) dénombre cinq opérations ou manipulations déployées dans les épreuves dont il brosse le portrait. Les tâches de jugement exigent des participant·es une prise de décision sur la relation sémantique entre deux mots. Certaines tâches demandent de segmenter des mots en plusieurs morphèmes, ou de mélanger (*blending*) des

morphèmes pour former des mots multimorphémiques, reprenant des opérations que nous retrouvons dans les tâches de dérivation et de décomposition. Les tâches d'analogie demandent de compléter une série de mots en imitant la structure dans une première série de mots (ex. *fleur : fleuriste, garage :* _____). L'identification d'affixes dans des pseudomots correspond pour Apel (2014) à un autre type d'opération. Dans le cadre des tâches d'identification d'affixes, une liste de pseudomots est fournie, et on demande aux participant·e·s d'identifier les suffixes qu'ils parviennent à repérer (Apel *et al.*, 2013).

La typologie des différentes opérations des tâches de conscience morphologique relevées dans le travail d'Apel (2014) n'est pas reprise intégralement dans la présente recherche. Nous avons plutôt cherché à reconceptualiser ces tâches de façon binaire : ajout et suppression. Ces deux opérations fondamentales correspondent à ce qui est demandé aux participant·es de faire pour réussir la tâche (Bernstein *et al.*, 2020; Tong *et al.*, 2014, p. 23). Elles sont ainsi susceptibles de concourir, en tant que variables latentes, à la réussite des tâches de conscience morphologique (Bourdages et Foucambert, 2018), et même des performances en compréhension.

2.2.2.4 Analyse et niveaux de contrôle

Carlisle (1995); (Carlisle et Nomanbhoy, 1993) est l'une des rares auteur·es à avoir conceptualisé les tâches de conscience morphologique à l'aune des modèles du

développement des habiletés métalinguistiques de Karmiloff-Smith (1986) et Gombert (1990). Dans ce cadre, les épreuves qui demandent une compréhension de morphèmes se distinguent fondamentalement de celles qui en nécessitent la production (Casalis *et al.*, 2002; Deacon *et al.*, 2008). Dans les épreuves de réception, par exemple lorsqu'on demande à l'élève de choisir, parmi quelques formes dérivées (flexionnelle ou dérivationnelle), celle qui convient le mieux, le niveau de contrôle est dit implicite (Carlisle, 1995; Deacon *et al.*, 2008). Ce type de tâche se différencie d'une situation de production, dans laquelle l'enfant est amené à donner une forme dérivée lorsqu'on lui fournit une base (Berko, 1958; Carlisle, 1995), ce qui correspond à un niveau de contrôle explicite (Duncan *et al.*, 2009).

Les tâches de « jugement » évaluent ainsi des habiletés plus implicites (ou épilinguistiques), car la demande cognitive pour réussir ce genre de tâche semble proche d'une connaissance fonctionnelle, implicite, de la langue (Duncan *et al.*, 2009; Goodwin *et al.*, 2017) – ou un traitement morphologique *tacite,* selon la terminologie de Goodwin *et al.* (2017) –, tandis que les tâches de « production », explicites – ou relatives à une analyse morphologique *stratégique* (Goodwin *et al.*, 2017), mesurent l'habileté à « analyser et manipuler » la structure morphémique des mots (Carlisle, 1995 p. 205).

Parmi les travaux qui se sont penchés sur la dimensionnalité de la conscience morphologique ou sur les facteurs de variation permettant de mieux circonscrire leur

mesure, nous n'avons pas retrouvé de trace évidente de la notion de contrôle attentionnel. Pourtant, il nous semble que certains des outils de mesure de la conscience morphologique impliquent une forme de contrôle de l'attention et d'inhibition des éléments non pertinents à la résolution du problème. Dans ce cadre, les épreuves impliquant des pseudomots demandent un niveau de contrôle élevé par rapport aux autres épreuves, et occupent ainsi une place sur le continuum. Tighe et Schatschneider (2015) ont effectivement montré qu'un modèle bidimensionnel traitant les tâches construites avec de vrais mots et celles construites avec des pseudomots comme deux facteurs latents distincts, mais corrélés, représentait mieux la variation dans leurs données.

La figure 2.1 correspond à une représentation graphique des différentes tâches de conscience morphologique et des demandes (spéculatives) qu'elles mobilisent en matière de contrôle et d'analyse, d'après le modèle de Bialystok (2001a).

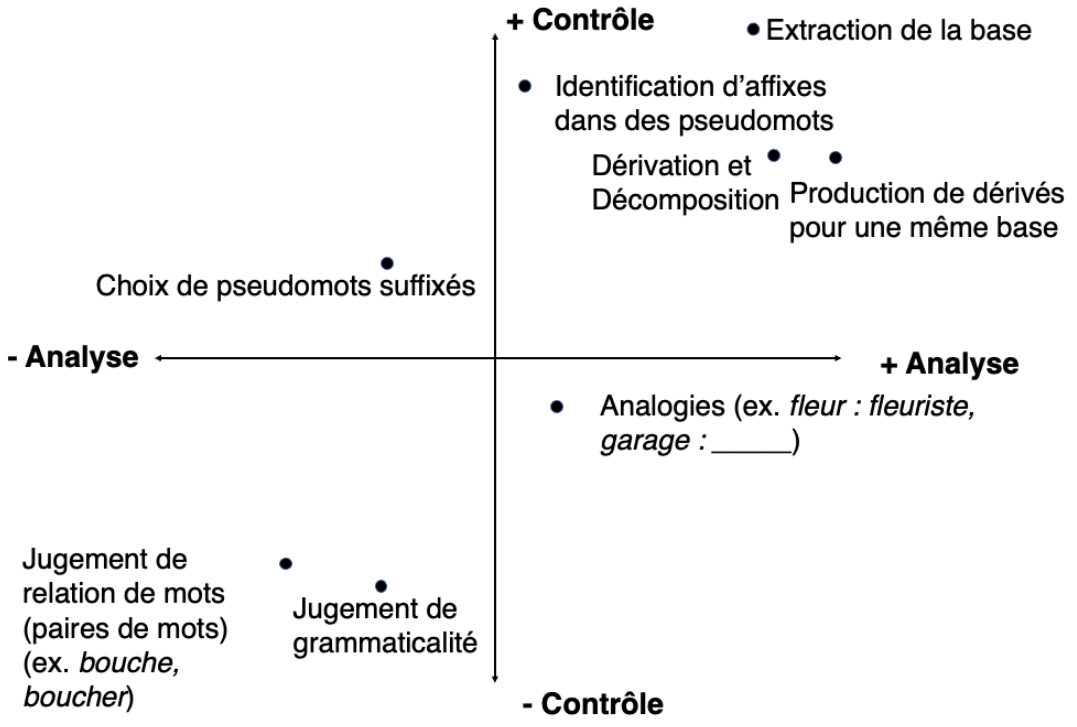


Figure 2.1 Proposition de représentation graphique des tâches mesurant la conscience morphologique, en fonction de leurs demandes en termes d'analyse et de contrôle, telles que définies par Bialystok (2001a)

2.3 Lien entre conscience morphologique et compréhension en lecture

La conscience morphologique est reconnue comme un prédicteur important et précoce des compétences de lecture des mots (Apel *et al.*, 2013; Bernstein *et al.*, 2020; Carlisle et Nomanbhoy, 1993 ; Casalis et Louis-Alexandre, 2000; Colé *et al.*, 2004; Kirby *et al.*, 2011; Nagy *et al.*, 2006) ainsi que des performances en compréhension en lecture chez les enfants de différents niveaux scolaires (Carlisle,

2000; Casalis *et al.*, 2002; Deacon *et al.*, 2014; Foorman *et al.*, 2018; Guo *et al.*, 2011; Kieffer et Lesaux, 2012; Kim *et al.*, 2019; Law *et al.*, 2015; Levesque *et al.*, 2017; Nagy *et al.*, 2003; Pan *et al.*, 2015; Simpson *et al.*, 2019; Wilson-Fowler et Apel, 2015) et chez les adultes (Tighe, 2015; Tighe et Binder, 2015; Tighe et Schatschneider, 2016; To *et al.*, 2016).

La relation entre la conscience morphologique et les performances en lecture est complexe dans la mesure où les morphèmes, a contrario des phonèmes, ont un sens, une structure, ainsi que des propriétés syntaxiques, phonologiques et relationnelles (Casalis et Louis-Alexandre, 2000; Mahony *et al.*, 2000). La conscience morphologique allie donc plusieurs aspects de la représentation linguistique et est, en ce sens, peut-être plus à même de donner un indice plus général des habiletés métalinguistiques que les consciences phonologique ou syntaxique prises isolément (Carlisle, 1995; Mahony *et al.*, 2000). La recherche dans le domaine de la lecture ne s'est pourtant intéressée que récemment au rôle de la conscience morphologique dans son apprentissage, l'accent ayant été mis sur la conscience phonologique et son influence sur la reconnaissance des mots écrits et la correspondance graphème-phonème (Desrochers *et al.*, 2009; Perfetti, 1985; Tunmer, 1989a). Certes, le code de l'écrit du français est alphabétique ; mais il est aussi morphémique.

Il est généralement considéré que la contribution de la conscience morphologique à la compréhension en lecture est unique, ou directe (Goodwin *et al.*, 2021), mais elle

est également indirecte, via le vocabulaire et la syntaxe (Levesque *et al.*, 2020; Levesque *et al.*, 2017), étant donné la propension des affixes à assumer des rôles sémantiques et syntaxiques. Ainsi, la contribution de la conscience morphologique à la compréhension en lecture des enfants va au-delà des compétences liées à la connaissance du vocabulaire (Mahony *et al.*, 2000; Nagy *et al.*, 2003; Singson *et al.*, 2000), la conscience phonologique (Carlisle, 1995; Carlisle et Nomanbhoy, 1993; Deacon et Kirby, 2004; Nagy *et al.*, 2003) et le QI verbal et non verbal (Deacon et Kirby, 2004; Kirby *et al.*, 2011). Le vocabulaire et la lecture des mots ont ainsi été identifiés comme des mécanismes clés qui sous-tendent les liens indirects entre la conscience morphologique et la compréhension de la lecture (Levesque *et al.*, 2017; Nagy *et al.*, 2006).

2.3.1 Sensibilité au traitement sémantique

L'habileté à repérer et à opérationnaliser les plus petites unités de sens de la langue est depuis une trentaine d'années réputée contribuer aux performances en lecture (Carlisle, 2000; Goodwin *et al.*, 2021; Kieffer *et al.*, 2016; Kuo et Anderson, 2006; Manolitsis *et al.*, 2017) et en en écriture (Green *et al.*, 2003). Par ailleurs, la contribution de la conscience morphologique à la compréhension augmente au fur et à mesure que s'amenuise celle de la conscience phonologique (Mahony *et al.*, 2000; Singson *et al.*, 2000). De la même façon, Carlisle (2000) observe que la conscience morphologique explique 43 % de la variance en compréhension écrite à 8-9 ans, puis à 55 % à 10-11 ans. Une première hypothèse concernant le rôle, par médiation ou

non, de la conscience morphologique dans les performances en lecture, est que l'apprenti-lecteur ou l'apprentie-lectrice pourrait être plus sensible au traitement sémantique des mots à lire qu'au décodage graphophonologique (Colé *et al.*, 2004).

Apel *et al.* (2013) observent le même patron de résultats, avec des différences marquées en fonction de la tâche utilisée : une tâche mettant l'accent sur les relations de sens entre des mots morphologiquement apparentés (par rapport à des tâches d'identification d'affixes, notamment) est plus à même de prédire les habiletés en décodage et en compréhension en lecture des élèves de maternelle (5-6 ans), de première année (6-7 ans) et de deuxième année (7-8 ans). De plus, la conscience morphologique n'explique pas de variance unique au-delà de la conscience phonologique dans les analyses de régression effectuées chez des élèves de 1re année (6-7 ans).

La contribution de la conscience morphologique à la compréhension est aussi médiée par d'autres construits. Levesque *et al.* (2017) ont démontré expérimentalement que la conscience morphologique contribuait directement et indirectement, par le biais de la lecture de mots isolés et de l'analyse morphologique (choix de définitions), à la compréhension en lecture chez des élèves de 8-9 ans. Les recherches de Nagy *et al.* (2003) portant sur des élèves « à risque » de 2e et 4e année (7-8 ans et 9-10 ans) offrent un portrait de la situation plus confus : pour les élèves de 7-8 ans, la conscience morphologique prend une part unique de variance de la compréhension de la lecture

(au-delà de sa covariance partagée avec la conscience phonologique et orthographique et le vocabulaire oral), tandis que chez les élèves plus âgés, le même construit n'a pu expliquer de part significative de variance dans les variables dépendantes testées.

2.3.2 Calcul syntaxique

Une deuxième hypothèse pouvant être mise en avant concernant le rôle de la conscience morphologique dans les performances en lecture concerne la spécificité de l'écrit et le calcul syntaxique. Comme l'affirme Gombert (1992), la lecture est différente de la compréhension à l'oral dans la mesure où, à l'écrit, les indices contextuels sont, à toutes fins utiles, absents. Par conséquent, le lecteur a besoin de se référer à la structure linguistique, incluant les informations flexionnelles et dérivationnelles, pour comprendre les informations présentées dans le texte (Casalis et Louis-Alexandre, 2000). Le repérage des indices dérivationnels permet l'interprétation sémantique des mots lus, tandis que les flexions constituent des indices importants pour les calculs sémantique et syntaxique de la phrase et des liens entre ses agents (Kuo et Anderson, 2006).

2.4 Synthèse du cadre théorique sur la conscience morphologique

La revue de la littérature sur la conscience morphologique nous a permis d'établir une définition opératoire de ce construit, à l'instar de Gombert (1990) et de Carlisle (1995),

et en écho à la terminologie métalinguistique dont nous avons fait état au chapitre I, la conscience morphologique est définie comme l'habileté à identifier et manipuler de manière consciente la structure morphologique des mots (Colé *et al.*, 2004). Sur le plan de l'acquisition de la morphologie dérivationnelle, nous avons vu que les connaissances morphologiques se développent sur une longue période, présentant une croissance marquée entre six et douze ans, mais continuant à s'améliorer tout au long de l'adolescence et à l'âge adulte. La morphologie flexionnelle se développe plus tôt, avant même la lecture. Les liens entre la conscience morphologique et la compréhension en lecture ont été exposés. Nous avons également fait état des types de connaissances (relationnelle, distributionnelle, syntaxique et réceptive) qui sont développées chez les enfants lors de leur acquisition d'habiletés en conscience morphologique. Les données expérimentales que nous avons sur ces connaissances et sur le caractère prédictif de la conscience morphologique vis-à-vis de la compréhension sont tributaires des instruments de mesure dont les recherches font usage.

Nous avons donc voulu porter un regard neuf sur la mesure de la conscience morphologique. Le portrait des épreuves nous a permis de dégager d'importants points de divergence. D'une part, les tests de conscience morphologique ne mettent pas en œuvre les mêmes unités linguistiques. Sont généralement mobilisés des éléments tant syntaxiques que morphologiques, puisque les épreuves requièrent, dans les faits, de supprimer les affixes de mots et de pseudomots à l'intérieur de phrases, ce qui demande effectivement aux sujets de focaliser leur attention sur des

éléments syntaxiques (Bernstein *et al.*, 2020). Par ailleurs, l'unité morphologique en jeu dans les tâches n'est pas stable : le morphème peut être flexionnel ou dérivationnel, et représenter un préfixe ou un suffixe (Casalis *et al.*, 2002). Les études ne font généralement pas mention d'un contrôle serré des items, notamment en ce qui a trait à leur transparence sur les plans phonologique et orthographique ou pour leur fréquence[38].

D'autre part, les tâches présentent des disparités sur le plan de la modalité de présentation : certaines sont présentées oralement ou par écrit, et les réponses peuvent être recueillies dans l'une ou l'autre de ces modalités (Deacon *et al.*, 2008). Les tâches impliquent des opérations cognitives de suppression et d'addition. Les niveaux de contrôle qui sont déployés dans le cadre de ces épreuves sont également différents (Apel, 2014; Casalis *et al.*, 2002).

Il est donc raisonnable de penser que la pluralité des mesures utilisées dans les études portant sur la conscience morphologique pourrait expliquer, au moins en partie, les écarts observés dans la prédiction d'habiletés de haut niveau essentielles au succès scolaire, notamment les compétences en lecture et en écriture (Apel *et al.*, 2013; Goodwin *et al.*, 2017; Tighe et Schatschneider, 2016). Pour savoir quelle est la part de variance qui explique le mieux les performances en compréhension en lecture,

[38] Les exceptions à ce constat sont le mémoire de Allard (2008), notamment, et les études dans lesquelles la transparence sémantique fait intégralement partie des objectifs de recherche, par exemple chez Hwaszcz et Klimek-Jankowska (2017) ou encore Diependaele *et al.,* (2009).

construit qui fera l'objet du chapitre IV, il est impératif de mieux circonscrire et mesurer les éléments qui varient dans les items, et de mieux contrôler les éléments qu'on veut constants. Cet objectif guidera le choix des épreuves et des items de la présente recherche, qui seront présentés dans la partie 6.2.

Marcel se pencha un peu plus vers Duplessis. « Y s'en viennent-tu la chercher ? » Duplessis se remit aussitôt sur ses pattes. « Ah ! oui, c'est vrai… Viens, Marcel, c'est à soir que l'âme de ta grand-mère va s'envoler. »

– Michel Tremblay, *La duchesse et le roturier*

CHAPITRE III

LA CONSCIENCE SYNTAXIQUE

Nous avons consacré le chapitre 2 à une revue de la littérature sur la conscience morphologique à partir de points de vue théorique, développemental et méthodologique. Le présent chapitre, qui garde globalement la même structure que le précédent, sera dédié à l'étude d'un second important construit métalinguistique : la conscience syntaxique. D'abord, nous présenterons la conscience syntaxique et son développement (3.1). Puis, dans la section 3.2, nous réviserons les épreuves étant utilisées pour mesurer la conscience syntaxique et les dimensions sous-jacentes qu'elles engagent. En dernier lieu, nous brosserons le portrait des liens directs et indirects entre le construit et la compréhension en lecture (3.3), puis ferons une synthèse (3.4).

3.1 Représentations syntaxiques

La découverte des propriétés profondes du langage humain, conçu comme une faculté et comme une fonction cognitive humaine complexe, est fondée sur la description linguistique des langues et la révélation de traits partagés entre elles, analysés dès lors comme potentiellement généralisables et fondamentaux (Hyams et Orfitelli, 2015). La généralisation a en effet lieu lorsqu'un ensemble de règles peut être mieux représenté par une règle plus générale, ou lorsque l'analyse démontre que des éléments linguistiques sont traités de manière analogue (Bouchard, 2019). L'étude de la syntaxe, en particulier, a été investie de cette approche méthodologique et théorique, notamment sous l'impulsion des travaux de Chomsky (1955/1975, 1957) et de l'avènement de la notion de grammaire universelle (GU) (Le Ny, 2017). L'analyse linguistique de la syntaxe correspond à l'étude de l'architecture des syntagmes et de leurs fonctions, des propositions et des phrases (Fayol *et al.*, 1992).

Le caractère hiérarchique de la représentation des phrases est un aspect fondamental de la théorie syntaxique (Chomsky, 1965), qui se vérifie sur le plan neurophysiologique (Nelson *et al.*, 2017) [39], donnant ainsi une idée de la façon dont est organisée la cognition humaine. La position structurelle des mots dans les phrases

[39] Pendant la lecture, notamment, l'activité neuronale dans les zones de l'hémisphère gauche dédiées au langage augmente à chaque mot successif, mais diminue soudainement chaque fois que des mots peuvent être fusionnés dans une phrase. Nelson *et al.*, 2017 proposent que cette diminution constitue une marque neuronale de la fusion syntaxique (« *merge* »), une opération fondamentale réputée représenter les propriétés récursives du langage humain (Hauser *et al.*, 2002). Ces positions autour des propriétés récursives du langage sont très controversées, et la présente thèse ne prétend pas prendre position à propos de ces débats.

contribue à leur sens, au-delà de la signification des éléments lexicaux (Schönefeld, 2011).

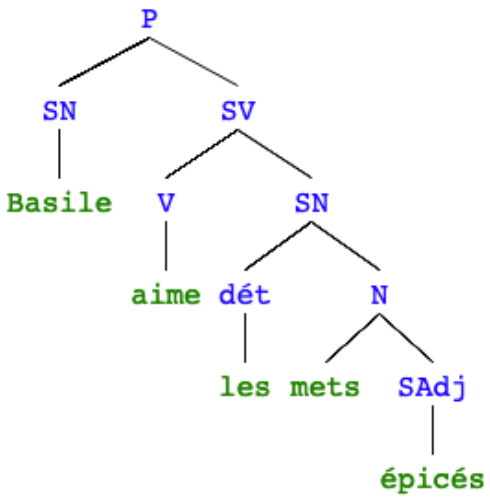


Figure 3.1 Représentation arborescente de la phrase « Basile aime les mets épicés »,où [P [SN Basile] [SV [V aime] [SN [dét[les]] [N[mets][SAdj[épicés]]]

Les arbres syntaxiques, comme celui présenté en 3.1, permettent de rendre compte de la structure des constituants de la phrase, et de représenter les relations hiérarchiques – de dépendance – qui les lient. Ils sont aussi censés représenter le caractère computationnel du traitement syntaxique, d'un point de vue cognitif, une position issue de l'explication computationnelle de l'esprit (*Computational Theory of Mind*, ou CTM). Dans le cadre de cette approche, les processus mentaux sont conçus comme des computations, comme c'est le cas dans la logique propositionnelle, et les

pensées sont des représentations mentales[40] qui possèdent une structure de type syntaxique et qui assument des rôles causaux qui dépendent de la forme syntaxique des représentations.

Le traitement syntaxique[41], malgré son apparente et évidente complexité, se produit hors de la conscience, et de façon automatique et implicite. Des recherches en imagerie cérébrale (*event-related brain potentials – ERPs*) ont en effet permis, par

[40] Cette idée fait écho à la théorie représentationnelle de l'esprit (ou RTM, mis pour *Representational Theory of Mind*)**,** dont les principales positions remontent au moins aux Grecs, et selon laquelle manifester des états mentaux correspond à avoir une représentation à l'esprit.

[41] L'acquisition de la syntaxe constitue un terreau fertile pour le débat entre nativistes et empiristes (*nature vs nurture*). La position générativiste nativiste argue notamment que certaines notions qui n'existent pas dans l'environnement sont exhibées par les enfants, ce qui correspond à une version de l'argument de la pauvreté du stimulus (Hauser *et al.*, 2002). Certaines recherches constructivistes empiristes dans le champ de l'acquisition tentent plutôt de démontrer une acquisition du langage purement basée sur l'input de l'environnement linguistique et le calcul statistique de correspondances (voir notamment les travaux de Saffran (2003) au sujet des régularités et de l'apprentissage du langage par des moyens statistiques, et Tomasello (2000) pour une théorie de l'acquisition basée sur l'usage). Elles sont les fruits des plus récentes avancées dans le domaine de l'apprentissage statistique de la syntaxe, notamment grâce à l'essor des mécanismes bayésiens de l'apprentissage. Nous nous garderons de tirer des conclusions fortes sur l'acquisition du langage. Le processus d'acquisition de la syntaxe est quant à lui très complexe : il implique l'apprentissage de l'ordre des mots propre à la langue (SVO, SOV), l'apprentissage de règles (passives, interrogatives, flexions casuelles), l'apprentissage de relation entre les constructions (items à polarité négative, inclusion, exclusion, etc.). Si certains sous-aspects du langage semblent être globalement bien prédits par les statistiques, celles-ci ne semblent pas suffire à expliquer l'acquisition du langage.

D'autres phénomènes d'acquisition pourraient possiblement être mieux circonscrits par des mécanismes computationnels symboliques que par un raisonnement probabiliste et logique. Il y a donc probablement des raisons computationnelles de penser que certaines parties du raisonnement, la prise de décisions et le langage ne se font pas par l'entremise du raisonnement bayésien, probabiliste et logique. Ces considérations s'éloignent des objectifs de la présente thèse. Nous tenons simplement pour acquis que certains éléments syntaxiques de la langue constituent des représentations auxquelles les élèves peuvent faire appel. Les mécanismes à la source de la compétence linguistique ne sont pas ici ce qui nous intéresse.

exemple, de mettre en lumière le fait que des erreurs, ou agrammaticalités ou violations grammaticales, produisent des réactions neuronales qui se situent hors de la conscience, c'est-à-dire qu'elles ne sont pas détectées consciemment (Batterink et Neville, 2013). En effet, la structure profonde des phrases et les marques transformationnelles font partie de la vie mentale inconsciente des locutrices et locuteurs (Friedmann et Reznick, 2021; Hyams et Orfitelli, 2015). L'intuition linguistique que les locutrices et locuteurs d'une même langue possèdent, ou développent en tout cas tôt dans le processus acquisitionnel (Friedmann et Reznick, 2021; Hyams et Orfitelli, 2015; Kail, 2000) leur permet de procéder à des jugements de grammaticalité ou d'acceptabilité syntaxique. Cette intuition correspond pour Chomsky (1965) à une manifestation de la compétence linguistique, soit un ensemble de connaissances et de règles partagées par les locutrices et locuteurs d'une même langue et d'une réalité mentale sous-jacente à la performance linguistique, responsable de tous les aspects de l'utilisation de la langue en production et en compréhension (Hyams et Orfitelli, 2015).

Toutefois, certaines unités linguistiques abstraites sont réputées pouvoir être objectivées, conceptualisées, ce à quoi s'emploient d'ailleurs les didacticien·nes qui travaillent sur la grammaire scolaire et l'utilisation du métalangage en classe[42]. Kail (2000, p. 14) affirme que « [d]ans sa conquête de la grammaire, l'enfant doit résoudre deux tâches essentielles : d'une part, isoler les éléments grammaticaux de la langue

[42] Citons, entre autres, Chartrand (2012, 2016) et Bulea Bronckart et Gagnon (2017).

et les règles qui président à leur ordre et, d'autre part, déterminer le sens de ces structures. » Les catégories grammaticales sont le niveau d'analyse syntaxique basique et peuvent être conceptualisées, objectivées. Pour certains auteurs, elles existent comme représentations spécifiques au langage (Pinker et Jackendoff, 2005). Pour ce qui est des groupes et de leurs fonctions sémantiques, Braine (1992) propose que des habiletés cognitives générales de *bootstrapping* sont suffisantes pour que même de jeunes enfants soient en mesure de distinguer un SN à fonction argumentale d'un SV à fonction prédicative[43] .

Il nous semble ainsi important de rappeler une propriété universelle de la syntaxe et du langage, comme envisagé, entre autres, par la grammaire générative, qui guide la plupart des entreprises cognitivistes à les comprendre : l'idée selon laquelle le langage et sa structure fonctionnent grâce à des éléments linguistiques que l'on peut conceptualiser – phonème, morphème, syntagme – qui peuvent être appliqués à toute langue (Kail, 2000; Ravid et Tolchinsky, 2002).

Les questions entourant la nature et le traitement des représentations syntaxiques dans la production et la compréhension du langage sont finalement assez éloignées

[43] Il faut tout de même rappeler que l'idée selon laquelle les locutrices et locuteurs ont acquis et possèdent des représentations des catégories syntaxiques abstraites de leur langue, notamment, est disputée. Ambridge (2017) adopte la position selon laquelle les catégories grammaticales sont « illusoires ». Les mots seraient emmagasinés en mémoire avec leurs probabilités combinatoires, sans que des représentations abstraites des catégories « verbe » ou « nom », par exemple, ne soient nécessaires.

des objectifs de la présente thèse, l'idée générale étant que ces représentations peuvent être manipulées et délibérément amenées à la conscience (Karmiloff-Smith, 1986; Ravid et Tolchinsky, 2002).

3.1.1 Conscience syntaxique

La conscience syntaxique[44] correspond à l'habileté à reconnaitre et à manipuler les structures syntaxiques d'une langue (Pratt *et al.*, 1984), et à contrôler intentionnellement les règles grammaticales de celle-ci (Gombert, 1990, p. 59). Nous l'avons mentionné préalablement : l'intérêt des chercheur·euses pour ce construit est motivé par sa qualité de prédicteur des habiletés de haut niveau (Gombert *et al.*, 1994; Kieffer *et al.*, 2016), et ce, tant en L1 qu'en L2. En effet, il est généralement reconnu que les habiletés métasyntaxiques facilitent l'analyse de phrases plus ambigües et la reconnaissance des mots pendant la lecture. Les habiletés métasyntaxiques se distinguent des habiletés syntaxiques ou épisyntaxiques, qui elles, mobilisent le traitement procédural implicite des structures syntaxiques (Simard, D. *et al.*, 2017).

[44] Dans le cadre de notre thèse, la « conscience syntaxique » (*syntactic awareness*) est considérée comme un synonyme des « habiletés métasyntaxiques » (*metasyntactic ability*), au même titre que sont interchangeables les notions d'« habiletés métalinguistiques » et de « conscience métalinguistique », comme nous l'avons mentionné dans la section 1.3. Le critère qui guide ce choix terminologique est méthodologique : ce qui définit et caractérise chacun de ces concepts est le matériel expérimental servant à les mesurer. Le terme « conscience syntaxique », qui fait l'objet du présent chapitre, a par ailleurs l'avantage de présenter un parallèle évident avec la « conscience morphologique » (*morphological awareness*), concept qui n'a, à notre connaissance, pas d'équivalent : « habiletés métamorphologiques » est inusité.

Deux différents types d'habiletés sont mis en œuvre dans les tâches cherchant à mesurer les habiletés métasyntaxiques, ou conscience métasyntaxique : d'une part, l'habileté à réfléchir sur la syntaxe, avec verbalisation de connaissances syntaxiques, qu'elles soient formelles ou informelles, et d'autre part, l'habileté à manipuler de façon délibérée les structures syntaxiques, sans verbalisation.

3.1.2 Développement métasyntaxique

L'âge d'apparition des habiletés métasyntaxiques est relatif à la définition adoptée, et tributaire des mesures utilisées. Ainsi, les autocorrections auxquelles procède l'enfant dès l'âge de deux ans peuvent être assimilées à une manifestation d'habiletés métalinguistiques (Gombert, 1990; Morgenstern *et al.*, 2014). Les autocorrections d'ordre syntaxique sont cependant extrêmement rares : elles représentent seulement 5% du total, la majorité d'entre elles étant essentiellement dédiées aux corrections d'ordres sémantique et lexical (Gombert, 1988). de Villiers et de Villiers (1972) rapportent que quelques enfants plus avancés parmi leur échantillon de sujets de 2-3 ans étaient en mesure de juger et de corriger de manière appropriée l'ordre des mots inversés dans des phrases impératives simples. En effet, dans des tâches de jugement impliquant des phrases asyntaxiques et asémantiques, les erreurs d'ordre sémantique sont souvent repérées et corrigées, mais les agrammaticalités ne sont repérées que par une poignée d'enfants, et rarement corrigées (de Villiers et de Villiers (1972) cité dans Gombert (1988)). L'étude de Pratt *et al.* (1984) sur des enfants anglophones de 5-6 ans, basée sur un protocole de corrections d'erreurs

morphologiques et de remise en ordre de mots, a aussi révélé cette tendance des sujets à fonder leurs jugements sur des considérations sémantiques plutôt que syntaxiques, dans le cas d'items de conscience métasyntaxiques. Les enfants les plus jeunes du corpus étaient particulièrement sujets à baser leur jugement sur l'acceptabilité sémantique des phrases.

Gombert (1990) rapporte que c'est vers l'âge de 6 ou 7 ans que les habiletés métasyntaxiques, comprises comme les habiletés à « focaliser l'attention des sujets sur des considérations de forme plutôt que de contenu » (Gombert, 1990, p. 80), se manifestent. Notons qu'il s'agit aussi de l'âge où commence la scolarité primaire dans plusieurs pays.

Les habiletés métasyntaxiques se développent vraisemblablement jusqu'à tard dans l'adolescence. Gaux et Gombert (1999b) rapportent que les tâches impliquant un contrôle et une explicitation, notamment la troisième version d'une tâche de réplication de l'erreur impliquant de porter volontairement son attention sur la syntaxe, induisent des résultats faibles dans leur échantillon de sujets âgés de 11 à 14 ans. L'étude de Marcotte (2021) révèle que les mesures d'identification de l'erreur et de remise en ordre de mots ont induit des scores plafond dans un échantillon adulte L1 (18 à 35 ans), mais une variation importante pour les résultats à la réplication de l'erreur (x = 86,1, σ = 12).

Un regard neuf porté sur les tâches de conscience syntaxique est un prélude à la tentative d'y voir plus clair dans les liens qui unissent habiletés métasyntaxiques et compréhension. La section suivante porte donc spécifiquement sur la mesure des habiletés métasyntaxiques.

3.2 La mesure des habiletés métasyntaxiques

Les sous-sections suivantes sont consacrées à l'établissement d'une typologie des tâches utilisées pour mesurer la conscience syntaxique.

3.2.1 Éléments convergents

Les instruments de mesure élaborés dans le cadre des recherches expérimentales en psycholinguistique et en acquisition du langage en langue première et seconde cherchent à identifier les indices de contrôle et de réflexion sur les unités syntaxiques chez les participant·es. Les tâches de conscience syntaxique font fonctionner des opérations cognitives, appliquées à des unités linguistiques (syntaxiques) qui sont mesurables : répétition de clauses, jugement de grammaticalité, identification et correction d'erreurs, etc.

Ces tâches impliquent de manière générale une analyse au niveau de *phrases isolées* ou, plus rarement, du mot[45], puisque certaines recherches assimilent les épreuves impliquant des morphèmes dérivationnels à un traitement syntaxique[46] (Simard, D. *et al.*, 2017). Par ailleurs, elles cherchent à mobiliser l'attention du participant sur la forme des énoncés, et non pas sur leur contenu sémantique, qui ne doit pas être saillant ni rendre difficile l'analyse syntaxique des items (Simard, D. *et al.*, 2017)[47].

Les objectifs de recherche des travaux en psycholinguistique et en acquisition du langage, de même que leurs corolaires méthodologiques, ont pour conséquence de générer des *épreuves non écologiques* : le dispositif de collecte de données assure une mesure rapide, efficace, ciblée. Le matériel expérimental est calculé, soupesé, afin de que les facteurs de variation soient le mieux possible contrôlés.

Le *temps* fait souvent partie de la mesure : s'il ne constitue pas nécessairement une variable, il est en tout cas souvent pris en compte. Un temps de traitement réduit

[45] Contrairement à certaines positions didactiques par exemple, dans lesquelles les analyses sont parfois situées dans le texte, même si les unités d'analyse en elles-mêmes peuvent être phrastiques ou lexicales.
[46] Cette position est bien sûr tout à fait défendable, et rejoint les positions génératives, pour lesquelles d'une part, l'étude de la structure interne des mots est analogue à la structure interne de phrases, et d'autre part, les morphèmes servent de nœuds terminaux dans les phrases, prouvant leur caractère syntaxique (Anderson, S.R., 2016). Toujours est-il que la recherche en psychologie cognitive a historiquement traité de la syntaxe et de la morphologie comme deux objets linguistiques distincts, dont les habiletés afférentes sont mesurables avec des outils différents, le premier portant sur la structure des phrases, et le second sur celle des mots (Kuo et Anderson, R.C., 2006).
[47] Les positions respectives et les influences réciproques des modules sémantique et syntaxique constituent un thème assez controversé en psychologie cognitive. Ce débat n'étant pas au cœur de nos réflexions, il ne sera pas traité dans le cadre de cette thèse.

suggère l'automatisation des traitements, et une intégration procédurale maximale si celle-ci est rapide[48].

La *modalité de passation* des épreuves de conscience syntaxique est très souvent orale. Ceci a pour effet de réduire l'influence du décodage sur les résultats. En revanche, cette technique peut occasionner une surcharge de la mémoire de travail (Brimo *et al.*, 2015; Nation et Snowling, 2000).

Un autre point de convergence important dans les travaux sur la conscience syntaxique concerne le *statut de l'erreur*. Celle-ci est mise à contribution pour la mesure du construit; de l'erreur, on infère l'intentionnalité du processus. Trois erreurs type semblent mobilisées dans les items répertoriés dans le matériel expérimental des épreuves, soit, comme en 1), un élément mal placé, comme en 2), un élément absent, et comme en 3), un élément erroné (Foucambert et Redmond, 2018; Marcotte, 2021).

1) * ***Des fleurs*** *la femme apporte.*
2) ** Le copain ___ ma mère est gentil.*

[48] Sur la recherche psycholinguistique et son rapport au temps, nous nous permettons de citer Le Ny (2017, p. 160) : « L'intérêt théorique de toutes ces mesures de « temps de... » tient à ce qu'elles constituent une expression mesurable d'une réalité plus profonde, le « temps de traitement » qui, dans l'esprit/cerveau des participants, correspond à l'exécution de la tâche considérée. Plusieurs décennies de recherche montrent deux choses : 1. Ces temps sont statistiquement très stables dans des conditions bien fixées ; 2. Ils sont statistiquement très variables en fonction des conditions. L'exploitation de ces deux propriétés fait d'eux des indicateurs précieux pour l'analyse fine du fonctionnement mental. Des modèles plus ou moins complexes, en fait de plus en plus sophistiqués, sont aujourd'hui proposés et testés pour en rendre compte ».

3) * *Le chat **dont** ta sœur aime est doux.*

Malgré ces éléments convergents, et à l'image de la diversité des tâches de conscience morphologique, problématisée dans la section 2.2, les épreuves qui mesurent les habiletés métasyntaxiques mettent en œuvre des habiletés latentes extrêmement diverses qui impliquent différents processus ou niveaux de contrôle, notamment sur un continuum allant des connaissances implicites aux connaissances explicites (Bialystok, 1988a, 2001a; Gaux et Gombert, 1999a; Simard, D. *et al.*, 2017). Les tâches donnent aussi à voir les distinctions conceptuelles entre les habiletés de haut et de bas niveaux, abordées dans la section 3.1.1. Les opérations mentales qui sont mobilisées dans les épreuves sont nombreuses et variées, celles-ci n'étant que rarement « pures » dans les opérations qu'elles demandent aux sujets (Gaux et Gombert, 1999b).

Nous ferons état de ces facteurs de variation par une taxonomie des tâches utilisées pour mesurer la conscience syntaxique, ou habiletés métasyntaxiques, dans les travaux portant sur ce construit[49].

[49] Cette taxonomie fait l'objet d'une discussion poussée avec visées comparatistes dans Bourdages *et al.* (2021). Certaines parties du texte ont été reprises intégralement dans le présent chapitre.

3.2.2 Description des épreuves métasyntaxiques

Nous passons ici en revue les tâches permettant de mesurer la conscience syntaxique au regard de leur caractère plus ou moins véritablement métalinguistique, en prenant comme jauge la nomenclature de Bialystok et Ryan (1985) et Bialystok (2001a), soit les habiletés d'analyse et l'exercice d'un contrôle sur les unités syntaxiques.

Parmi les tâches couramment utilisées utilisées pour mesurer les habiletés métasyntaxiques, on retrouve notamment la tâche de *jugement de grammaticalité*, qui consiste à demander aux participants de juger de la grammaticalité de certaines phrases (Bialystok, 1986; Cohen, 2011; Gaux et Gombert, 1999b; Gombert *et al.*, 1994; Hermanto *et al.*, 2012; Plaza et Cohen, 2003). À l'instar de l'épreuve de *préférence grammaticale*, la tâche de jugement est utilisée pour mesurer le niveau d'aisance linguistique, et consiste alors en une mesure des connaissances implicites et de connaissances de l'ordre de l'intuition linguistique, si les phrases utilisées sont sémantiquement acceptables (Bialystok, 1986; Simard, D. *et al.*, 2017). Étant donné que des habiletés syntaxiques seules suffisent à y répondre (Simard, D. *et al.*, 2017), leur caractère métalinguistique est ainsi contestable et débattu (Cohen, 2011).

La tâche de *localisation de l'erreur* (Gaux et Gombert, 1999a, 1999b; Gombert *et al.*, 1994) permet de voir si le sujet est en mesure d'identifier l'erreur dans une phrase. La *correction de l'erreur* (Bowey, 1986a; Cain, 2007; Diaz, 1985; Gaux et Gombert, 1999b; Gombert *et al.*, 1994), épreuve proche de la précédente, dans laquelle une

forme grammaticale correcte doit être substituée à une forme erronée, ne relève pas des habiletés métasyntaxiques, pour plusieurs auteur·es (Gaux et Gombert, 1999a, 1999b; Simard, D. *et al.*, 2017), puisque le recours à des constructions internalisées et automatisées est suffisant pour la réussir. Ainsi, ce sont les habiletés syntaxiques, et non métasyntaxiques, qui sont susceptibles d'être mobilisées pour répondre à la tâche.

La tâche de remise en ordre de phrases (*syntactic word-order correction task*) (Blakmore et Pratt, 1997; Brimo *et al.*, 2015; Cain, 2007; Deacon et Kieffer, 2018; Demont et Gombert, 1996; Gaux et Gombert, 1999a; Pratt *et al.*, 1984) implique le plus souvent des phrases dont les mots ont été aléatoirement mélangés. Le sujet doit remettre en ordre les différents items lexicaux afin d'obtenir une phrase qui ait du sens.

Dans la tâche de *répétition de l'erreur (ou* répétition de phrases agrammaticales (Bowey, 1986a; Bowey et Patel, 1988; Cain, 2007; Cohen, 2011; Foorman *et al.*, 2018; Gaux et Gombert, 1999b; Simard, D. *et al.*, 2012b), une phrase agrammaticale est présentée à l'oral, puis le sujet doit la répéter verbatim. Le fait de réprimer la correction automatique de la phrase agrammaticale constitue une manifestation d'un certain contrôle sur les éléments syntaxiques (Simard, D. *et al.*, 2011), ce qui permet de statuer sur son caractère « méta ». On tient pour acquis que le sujet doit focaliser son attention sur l'erreur, de façon intentionnelle, et réprimer l'envie plus intuitive, qui relève de connaissances implicites, de la corriger (Gaux et Gombert, 1999a).

Les habiletés métasyntaxiques sont aussi évaluées à l'aide de protocoles de *réplication d'erreurs*, un devis expérimental initialement paru dans (Gombert *et al.*, 1994). Comme mesure des habiletés métasyntaxiques, la réplication d'erreurs a essaimé dans les pays francophones dans le champ de l'acquisition en L2 (Armand, 2000; Simard, D. *et al.*, 2011; Simard, D. *et al.*, 2012b) et dans les travaux sur les habiletés métasyntaxiques en L1 (Gaux et Gombert, 1999a, 1999b). Cette tâche demeure cependant largement ignorée par les chercheur·es anglophones. Lors de la réalisation de cette tâche, le sujet doit écouter une phrase contenant une erreur grammaticale, comme « la fillette un film regarde » (Gaux et Gombert, 1999b), et reproduire la même erreur dans une autre phrase, celle-ci syntaxiquement parfaitement constituée. Cette épreuve suppose l'application de règles grammaticales et des capacités de manipulation d'éléments linguistiques, au même titre que les épreuves transformationnelles utilisées dans les épreuves de conscience phonologique, notamment (Gombert *et al.*, 1994). Pour Gaux et Gombert (1999b, p. 70), il s'agit de la seule tâche qui constitue une mesure certaine de la conscience métasyntaxique, puisque sa réussite demande de mettre réellement en œuvre des procédés de réflexion intentionnelle sur la syntaxe.

De plus rares mesures de la conscience syntaxique incluent la formulation de phrases (Kieffer *et al.*, 2016). Dans cette épreuve, les sujets doivent énoncer une phrase à

partir d'une image en utilisant 1 ou 2 mots-cibles présentés oralement par l'expérimentateur·trice.

D'après notre recension des écrits méthodologiques sur la conscience syntaxique, peu de place est accordée à la verbalisation et l'explicitation. Le lien entre explicite et explicitation est peu théorisé, ni, semble-t-il, vu comme nécessaire. Par exemple, on pourrait imaginer une tâche d'explicitation d'erreur impliquant une connaissance explicite de la règle grammaticale qui fait défaut (Gombert *et al.*, 1994), qu'elle soit formelle ou informelle (Simard, D. *et al.*, 2017).

3.2.3 Dimensionnalité de la conscience syntaxique

Nous l'avons mentionné : l'évaluation de la conscience métasyntaxique est effectuée à l'aide d'épreuves variées et contrastées, et par conséquent, les habiletés qu'elles sont censées mobiliser diffèrent aussi (Bialystok, 2001a; Simard, D. *et al.*, 2017). Ainsi, le même construit théorique est mesuré à l'aide d'épreuves foncièrement différentes. Souvent, les recherches sur la métasyntaxe et son lien avec la compréhension ne sont fondées que sur une ou deux épreuves permettant de l'évaluer. On retrouve notamment dans la littérature des épreuves de remise en ordre de mots, de jugements de grammaticalité, de correction d'erreurs, d'explicitation (formelle ou non) d'erreurs, de répétition d'erreurs ou encore de réplication d'erreurs. Nous inscrivant dans la continuation du mouvement méthodologique initié par Bialystok et Ryan (1985), nous tenterons de mieux définir en quoi les différents

instruments de conscience syntaxique divergent, afin de prendre en considération ces facteurs de variation dans nos analyses ultérieures.

La diversité avérée des tâches permettant de mesurer la conscience métalinguistique, et les habiletés métasyntaxiques notamment, et surtout la diversité des résultats exhibés par les sujets bilingues par rapport à leurs pairs monolingues, a motivé la démarche méthodologique de Bialystok et Ryan (1985) et de Bialystok (2001a).

Cette démarche méthodologique a permis de dégager deux dimensions, imaginées comme deux continuums sur lesquels les épreuves peuvent être projetées, et jugées les unes par rapport aux autres. Pour rappel, ces deux dimensions sont *l'analyse des connaissances linguistiques* et le *contrôle des traitements langagiers*. Nous nous intéresserons à la dimensionnalité des épreuves de conscience syntaxique sous cet angle méthodologique (en 3.2.3.1 et 3.2.3.2), puis nous procéderons à une reconceptualisation de ces épreuves, en imaginant d'autres types de continuums, d'autres dimensions orthogonales présidant à leur réalisation (en 3.2.3.3 et 3.2.3.4).

3.2.3.1 Niveaux de contrôle

Dans le chapitre I, nous avons présenté une série d'arguments pour montrer le caractère subjectif de la notion de « conscience » et pour justifier de l'abandonner dans notre travail sur typologie des tâches métalinguistiques. Nous lui avons préféré la notion de « contrôle », que nous réinvestissons dans notre revue méthodologique

de la conscience syntaxique.

Les épreuves utilisées pour mesurer les habiletés métasyntaxiques mobilisent des traitements ou niveaux de contrôle plus ou moins élevés. Certaines tâches métasyntaxiques mobilisant des capacités de contrôle attentionnel plus élevées peuvent exiger que les sujets concentrent leur attention sur certains éléments linguistiques, tout en ignorant ou en inhibant une compréhension du sens ou d'informations entrant en contradiction avec le but de la tâche et le construit visé (Cohen, 2011; Gaux et Gombert, 1999b; Simard, D. *et al.*, 2017). Ce niveau de contrôle attentionnel plus élevé, induit par l'inhibition d'éléments discordants et par l'intentionnalité dans la direction de l'attention, se traduit par une charge cognitive plus lourde (Cohen, 2011), comme c'est le cas pour la répétition de phrases agrammaticales, notamment.

La tâche de réplication de l'erreur est exemplaire de ce type de contrôle attentionnel. En effet, le participant doit reproduire, dans une phrase correcte, une erreur qu'il aurait préalablement identifiée dans une phrase incorrecte. Cette tâche est réputée mobiliser une réflexion intentionnelle sur la structure, et le moins possible impliquer de variables autres, comme des connaissances implicites sur la structure grammaticale ou le vocabulaire (Gaux et Gombert, 1999b).

3.2.3.2 Analyse des connaissances syntaxiques

Dans le modèle de Bialystok et Ryan (1985), l'axe qui représente l'analyse des connaissances linguistiques incarne cette habileté, mobilisée de manière inégale dans les différentes tâches, à se représenter des structures linguistiques de plus en plus explicites, fines et abstraites (Bialystok, 1986). Les épreuves de conscience syntaxique varient fortement sur ce plan, allant d'un traitement implicite des unités syntaxiques, comme l'identification d'erreurs, à un traitement explicite de celles-ci, l'explicitation culminant avec la verbalisation de règles grammaticales informelles ou formelles (Simard, D. *et al.*, 2017). Pour Gaux et Gombert (1999a), les tâches de répétition, de jugement, de correction et de localisation peuvent être réalisées de manière implicite (celles-ci ont induit des effets plafond dans leur échantillon de sujets âgés de 11 ans 5 mois à 14 ans 4 mois).

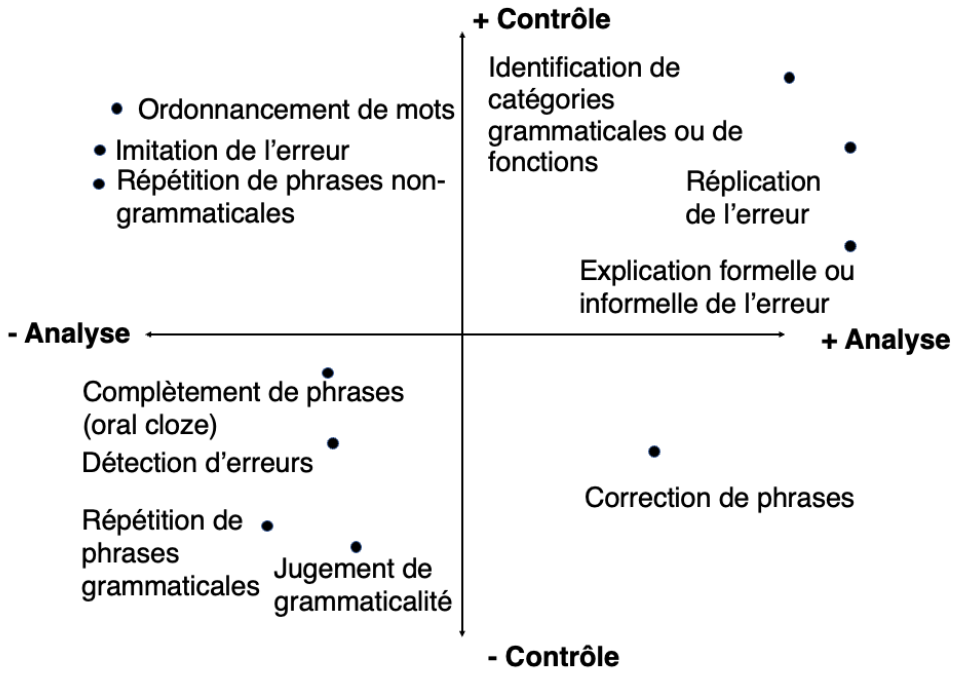


Figure 3.2 Le modèle de Bialystok (2001a) avec projection des tâches psycholinguistiques issues de Besse, A.-S. *et al.* (2010), Simard, D. *et al.* (2017), Cohen (2011) et Lesaux et Siegel (2003).

La figure 3.2 permet de représenter les niveaux de contrôle et d'analyse requis pour réussir les tâches métasyntaxiques, que nous avons répertoriées en 3.2.2. Les épreuves situées dans les valeurs négatives de l'axe des abscisses (x) demandent un niveau d'analyse faible, et dans les valeurs positives, un niveau plus élevé ; celles figurant dans les valeurs négatives de l'axe des ordonnées (y) mobilisent des niveaux de contrôle plus faibles, et dans les valeurs positives, des niveaux plus élevés. La place des épreuves sur le graphique, et donc les niveaux de contrôle et d'analyse qu'elles mettent en œuvre, a été évaluée en partie à partir des travaux et recommandations de Bialystok (2001a), Gaux et Gombert (1999a), Fortier (2013),

Simard, D. *et al.* (2017) et, dans une certaine mesure, Cohen (2011)[50], ainsi que de nos propres analyses.

Le modèle présenté à la figure 3.2 prévoit que la tâche de réplication de l'erreur mobilise plus d'habiletés de contrôle que d'analyse, alors que l'explication formelle ou informelle de l'erreur fait appel à plus d'habiletés d'analyse que de contrôle. La tâche d'explication de l'erreur demande véritablement une verbalisation des phénomènes grammaticaux en jeu et mobilise une série de processus, allant de la détection et l'analyse de l'erreur à l'utilisation de notions grammaticales. La détection d'erreurs, dans le quadrant en bas à gauche, fait plutôt appel à des connaissances implicites de la langue (Gaux et Gombert, 1999b). Le recours aux connaissances explicites de la structure grammaticale est requis pour répondre aux corrections d'erreurs et aux tâches de réordonnancement de phrases (Cohen, 2011).

3.2.3.3 Types de syntagmes (unités syntaxiques)

Au-delà des deux dimensions suggérées par Bialystok et Ryan (1985), cet effort de reconceptualisation des tâches psycholinguistiques permet d'imaginer plusieurs couples de dimensions orthogonales, pensées comme les extrémités de plusieurs séries de continuums, à l'image de ce que nous avons pu établir comme taxonomie

[50] Cohen (2011) a adopté la nomenclature de Bialystok (2001a) pour conceptualiser ses épreuves de conscience syntaxique, mais sans en exploiter le caractère continu des deux dimensions. L'auteure a plutôt opté pour une analyse binaire en fonction de la dimension qui était la plus saillante pour chaque épreuve.

dans notre chapitre II sur les tâches de conscience morphologique. Les éléments syntaxiques sont inégaux sur le plan de la charge cognitive qu'ils représentent, en fonction de leur niveau d'abstraction (voir Brimo *et al.* (2018); Simard, D. (2008)). Les items dans les épreuves varient sur le plan du type de phrases (actives, passives), de la longueur et de la complexité des phrases (Brimo *et al.*, 2018; Cohen, 2011). Les auteur·es ne font pas toujours état des unités syntaxiques mises en œuvre dans les épreuves, alors que la nature des items peut certainement avoir un effet sur les résultats.

Les erreurs sur lesquelles les sujets doivent focaliser leur attention peuvent être intrasyntagmatiques et porter sur des catégories grammaticales (ex. nom, déterminant, verbe, subordonnant, préposition) – comme dans *chat le regarde sa gamelle, où il y a inversion du déterminant et du nom dans le GN. Elles peuvent aussi porter sur des fonctions grammaticales, comme dans *le zèbre le lion dévore (Gaux et Gombert, 1999b). Dans cet exemple, le complément direct est situé en début de phrase et rend la phrase agrammaticale (SVO --> OSV). La phrase en (1) permet aussi d'exemplifier un problème maintes fois reconnu dans les épreuves de conscience syntaxique : celui de l'apport de la sémantique dans la mesure du construit. En effet, la focalisation délibérée sur la forme syntaxique n'est pas le seul facteur permettant de résoudre l'agrammaticalité de la phrase : le sens qu'on donne aux différents arguments a certainement son rôle à jouer (Gaux et Gombert, 1999b).

La mesure de la conscience syntaxique se fait aussi sur la base de phrases passives (Nation et Snowling, 2000), de phrases avec propositions relatives, de phrases mettant en œuvre des comparaisons (ex. **Bill is more smaller than Bob* (Bowey, 1986b))[51], de pronoms (ex.**I know what them are*. (Bowey, 1986b)), des règles d'accord (ex. **the baby are sick* (Cain, 2007)), des règles de morphologie flexionnelle (ex. **Yesterday, John learn his spellings* (Cain, 2007)), ou encore des prépositions (ex.**John gave the crayon for Mary* (Deacon et Kieffer, 2018)).

3.2.3.4 Opérations

Nous avons démontré dans la section 3.2.3.3 que les tâches de conscience syntaxique impliquent des unités linguistiques (syntaxiques) variées. Préalablement, dans la section 3.2.2, nous avons pu établir le fait que ces unités syntaxiques sont mises en œuvre dans de nombreux types d'opérations : répétition de phrases grammaticales et agrammaticales, jugement de grammaticalité, identification/localisation d'erreurs, correction d'erreurs, test de closure (*cloze task*), réplication de l'erreur (ou production analogue de phrases agrammaticales, réordonnancement de phrases (word ordering).

[51] Dans l'étude de Bowey (1986b), les phrases agrammaticales testées représentent des erreurs attestées dans le discours des enfants anglophones. Les travaux subséquents ont en général cherché à mettre en œuvre des agrammaticalités plus évidentes.

L'analyse plus fine des tâches mesurant la conscience syntaxique laisse entrevoir finalement que des opérations beaucoup plus atomiques concourent à la réussite des différentes tâches de conscience syntaxique. À l'image des épreuves de conscience morphologique, l'ajout et la suppression sont susceptibles d'être à la source d'une part de variation. Nous faisons l'hypothèse que ces deux opérations sont mobilisées dans des épreuves de déplacement – qu'il s'agisse d'éléments internes aux syntagmes (ex. *The teacher the story read to the children (Deacon et Kieffer, 2018)) ou de syntagmes entiers (ex.*Interested in music Mary wasn't. (Deacon et Kieffer, 2018)) – ou de remise en ordre des mots. Il en est de même pour les épreuves de remplacement.

3.3 Lien entre habiletés métasyntaxiques et lecture

Les liens entre la compréhension en lecture et les habiletés métasyntaxiques procèdent de l'idée générale selon laquelle la maitrise du code écrit requiert le développement d'habiletés spécifiques, non pas seulement en lien avec la modalité, mais au regard de la formalisation de la langue (Bowey, 2005; Gombert *et al.*, 1996), la lecture pouvant être décrite comme une tâche linguistique formelle. Ainsi, la tâche de l'apprentie lectrice ou apprenti lecteur, ou de la lectrice ou du lecteur, ne consiste pas seulement en le traitement perceptif de l'input visuel écrit, mais en la mise en place de mécanismes d'analyse explicite de la langue et d'un contrôle cognitif dont la complexité est plus élevée que ce qui doit être mobilisé pour le traitement de l'oral (Demont et Gombert, 1996; Perfetti *et al.*, 2005).

Le raisonnement théorique sous-jacent à cette entreprise de compréhension des habiletés métasyntaxiques comme prédicteurs de la compréhension en lecture est que les habiletés syntaxiques assurent une décomposition efficiente des phrases formées de plusieurs éléments ou groupes de mots (*chunks*) plus faciles à manier (Farnia et Geva, 2013), ce qui facilite la formation d'une représentation cohérente de l'ensemble du texte (Demont et Gombert, 1996; Nation et Snowling, 2000; Verhoeven et Perfetti, 2008). De bonnes habiletés métasyntaxiques sont réputées permettre la reconnaissance des mots dans leur contexte (Perfetti *et al.*, 2005) et l'analyse (*parsing*) des phrases complexes (Nation et Snowling, 2000) ou ambigües en l'absence d'indices prosodiques – puis la révision de cette analyse en cas d'échec (Bowey, 1986b).

Un débat fondamental dans la littérature sur le lien entre les habiletés métasyntaxiques comme variable indépendante et la compréhension comme variable dépendante est celui de la qualité du rôle prédictif que jouent les habiletés métasyntaxiques (Gombert *et al.*, 1994). Les études qui investiguent les liens entre les habiletés métasyntaxiques comme VI et la compréhension comme VD incluent généralement un ensemble de facteurs de contrôle – typiquement, la conscience phonologique (Demont et Gombert, 1996; Gombert *et al.*, 1994; Plaza et Cohen, 2003), les habiletés syntaxiques (Brimo *et al.*, 2015), le vocabulaire (Deacon et Kieffer, 2018) et certaines habiletés non verbales, comme le raisonnement (Deacon et Kieffer, 2018) – afin d'évaluer le poids des différentes variables et les liens de médiation les

unissant à la compréhension. Le rôle prédictif de la conscience syntaxique est tour à tour présenté comme direct (Brimo *et al.*, 2015; Deacon et Kieffer, 2018; Gombert *et al.*, 1994) ou indirect, c'est-à-dire médié par d'autres habiletés (Blakmore et Pratt, 1997; Brimo *et al.*, 2015; Cain, 2007; Perfetti *et al.*, 2005; Tunmer, 1989b, 1990).

Parmi les liens de médiation entre la conscience syntaxique et la compréhension en lecture ayant été théorisés, puis évalués expérimentalement, celui impliquant le décodage et la lecture de mots isolés, dans la mouvance du modèle simple de la lecture (*Simple View of Reading*) (Gough et Tunmer, 1986), est probablement le plus fécond. Les recherches de Tunmer (1989b, 1990), faites avec des sujets de 6-8 ans (grade 1-2), ont révélé des relations médiées par le décodage. Pour Tunmer (1990), les habiletés métasyntaxiques sont essentielles à la reconnaissance des mots écrits. Les enfants mobiliseraient leurs connaissances syntaxiques pour pallier une analyse phonologique erronée de mots opaques non familiers ou de mots homographiques. Cette position suggère que les connaissances phonologiques et les connaissances syntaxiques se développent en parallèle et concourent à l'apprentissage de la compréhension en lecture par la médiation du décodage et de la reconnaissance des mots écrits. Perfetti *et al.* (2005) propose un autre mécanisme expliquant le rôle indirect de la conscience syntaxique dans la compréhension en lecture en faisant l'hypothèse que la syntaxe fait partie d'une représentation lexicale de qualité.

D'autres études montrent que la relation entre la conscience syntaxique et la compréhension de la lecture va au-delà du contrôle de la lecture de mots, ou décodage graphophonologique, mesurée par des épreuves de lecture de listes de mots et de pseudomots (Bowey, 1986b; Brimo *et al.*, 2015; Deacon et Kieffer, 2018; Nation et Snowling, 2000). Gaux et Gombert (1999a) ont relevé des changements dans les contributions de R^2 dans les régressions multiples à ordre fixe de six épreuves de conscience syntaxique de l'ordre de 4,49% à 10,25% sur une mesure de la compréhension chez des sujets adolescents. En outre, Deacon et Kieffer (2018) ont mesuré une contribution directe dans leur échantillon de sujets 8-10 ans (*grade 3-4*), analysant que la conscience syntaxique est en mesure de prédire les performances en compréhension réalisées lorsque le décodage est bien en place, mais pas au début de l'apprentissage de la lecture. L'étude de Nation et Snowling (2000), dans laquelle des groupes de sujets ayant une bonne et une mauvaise compréhension de la lecture, appariés en fonction de l'âge, du décodage et des habiletés non verbales, a révélé que les enfants mauvais compreneurs possèdent des faiblesses générales dans le traitement de la langue. Ces recherches mettent en lumière un rôle direct causal de la conscience syntaxique sur la compréhension et laissent sous-entendre que la conscience syntaxique est un aspect des habiletés métalinguistiques qui sont mobilisées lors de la compréhension en lecture.

Par ailleurs, les études ayant révélé que la conscience syntaxique ne contribue pas à la compréhension de la lecture ont généralement été menées sur des échantillons

d'enfants plutôt jeunes (notamment Bowey et Patel (1988) avec des sujets de 7 ans et Cain (2007) avec des sujets de 8-10 ans). L'ensemble de ces résultats suggère un possible effet développemental sur les liens de causalité entre conscience syntaxique et compréhension. Il tend en effet à renforcer la position selon laquelle la conscience syntaxique contribue principalement aux processus d'intégration du texte, facilitant ainsi la compréhension de la lecture, et permettrait de prédire une corrélation plus forte entre la conscience syntaxique et la compréhension de la lecture qu'entre la conscience syntaxique et l'aptitude au décodage des mots (Nation et Snowling, 2000).

Dans l'étude de Gaux et Gombert (1999a), les analyses de régression ont révélé des contributions différenciées de la compréhension selon la tâche syntaxique, classées en fonction de leur niveau d'explicitation sur le plan de l'analyse des connaissances syntaxiques, l'un des axes dégagés par Bialystok. Plus les tâches de conscience syntaxique demandent une connaissance explicite de la syntaxe, donc moins elles sont implicites, plus leur contribution aux performances en compréhension est importante (Gombert *et al.*, 1994).

La conscience syntaxique dans plusieurs des études citées a été mesurée à l'aide de tâches de répétition et de correction d'erreurs ou d'ordre de mots (voir notamment (Blakmore et Pratt, 1997; Bowey, 1986b; Bowey et Patel, 1988; Cain, 2007; Deacon et Kieffer, 2018). Le caractère *méta* de ces deux tâches, c'est-à-dire les demandes qu'elles impliquent sur le plan de l'analyse explicite des structures syntaxiques et sur

le plan du contrôle cognitif, est assez controversé (Simard, D. *et al.*, 2017) (cf. section 3.2). Les mécanismes qui sous-tendent les liens de médiation et les liens directs et causaux entre les habiletés métasyntaxiques et la compréhension ne sont ainsi pas complètement élucidés, et doivent être étudiés empiriquement.

3.4 Synthèse du cadre théorique sur la conscience syntaxique

Cet examen des travaux portant sur les habiletés métasyntaxiques, ou la conscience syntaxique, au regard de la littérature portant sur les habiletés métalinguistiques exposées au chapitre I, nous a permis de déterminer une définition opératoire du construit. Il correspond à l'habileté à reconnaitre et à manipuler les structures syntaxiques d'une langue (Pratt *et al.*, 1984), et à en contrôler intentionnellement les règles grammaticales (Gombert, 1990, p. 59). L'intérêt pour les habiletés métasyntaxiques en langue première est fondé sur son caractère prédictif de la compréhension en lecture. En effet, elles sont réputées faciliter l'analyse de phrases ambigües et la reconnaissance lexicale pendant la lecture.

En ce qui a trait à la mesure de la conscience syntaxique, nous avons relevé plusieurs éléments de convergence :

- la modalité de passation des épreuves, qui est le plus souvent orale.
- le statut de l'erreur, qui permet d'inférer l'intentionnalité du processus.
- l'importance du temps de traitement.

La fragmentation des tâches qui mesurent les habiletés métasyntaxiques a permis de dégager des dimensions sur lesquelles elles se différencient. Elles sont inégales sur le plan de leurs demandes en matière d'analyse des connaissances et de contrôle des traitements langagiers. Les tâches diffèrent aussi sur le plan de l'unité : les erreurs portent sur des éléments syntaxiques variés. Elles se distinguent aussi quant aux opérations qu'elles mettent en œuvre : jugement, localisation, correction, correction de l'ordre des mots, réplication. À l'instar des épreuves de conscience morphologique, toutes ces formes de manipulations syntaxiques impliquent à tout coup deux formes d'opérations beaucoup plus élémentaires : l'ajout et la suppression.

La diversité des instruments de mesure utilisés dans les études portant sur la conscience syntaxique pourrait expliquer la variation observée dans la prédiction de la compréhension en lecture, qui fera l'objet du prochain chapitre.

Explanations of complex cognitive processes have too many degrees of freedom.

— Walter Kintsch, *Comprehension: A paradigm for cognition*

CHAPITRE IV

LA COMPRÉHENSION EN LECTURE

Nous avons présenté dans les chapitres précédents la conscience morphologique et la conscience syntaxique, deux construits étudiés en grande partie parce qu'ils prédisent la compréhension en lecture. Dans ce chapitre, nous ferons état de quelques remarques préliminaires sur le construit (4.1) et sur son caractère multicomponentiel (4.2). Nous nous pencherons sur la compréhension en lecture telle qu'elle est conçue dans la littérature en psychologie cognitive, c'est-à-dire en fonction des représentations (4.3) et des processus qu'elle met en œuvre (4.4), puis sur sa mesure (4.5).

4.1 La compréhension du discours : remarques préliminaires

La compréhension, dans son sens large, est un processus de première importance pour la psychologie cognitive et les sciences cognitives plus généralement (Graesser *et al.*, 1997). Pierre angulaire de la cognition – « *the backbone of cognition* » (McNamara et Magliano, 2009, p. 298) –, elle sert de base à d'autres processus éminemment cognitifs, comme le raisonnement, l'apprentissage, la formation de souvenirs, la catégorisation, la résolution de problèmes, la prise de décisions ou même les réponses affectives d'ordre esthétique face aux œuvres d'art (Millis et Larson, 2008). Rappelons ainsi d'entrée de jeu que la compréhension, qui consiste en la construction de plusieurs niveaux de représentations d'un phénomène (Graesser *et al.*, 1994), est un prérequis dans le traitement des informations rencontrées dans la vie de tout être humain. Ces informations sont de natures diverses; conversations, images, sons et autres inputs environnementaux variés participent à la compréhension.

La recherche sur la compréhension en psychologie cognitive s'est d'abord développée du côté du traitement du discours (Graesser *et al.*, 2003; Graesser *et al.*, 1997; Perfetti et Lesgold, 1977; van Dijk et Kintsch, 1983). Par commodité, mais aussi pour l'intérêt que représente la lecture pour qui s'intéresse à son enseignement et à son apprentissage, c'est la compréhension de textes écrits qui a été la plus investie par les chercheur·es.

Les termes « discours » et « texte » recouvrent également le discours oral (Bronckart, 2004; McNamara et Magliano, 2009). Les modèles de la compréhension du discours se veulent donc inclusifs, en embrassant les textes écrits et oraux. Leur portée va au-delà de la compréhension du langage; ces modèles s'appliquent au traitement cognitif en général, à tout ce qui peut être « compris ». Citons en exemple les travaux de Gernsbacher (1990), dont la position sur les processus en jeu dans la compréhension est explicitement *domain-general*.

Il faut ainsi concevoir que les performances en lecture sont également fonction, entre autres, des capacités fonctionnelles orales de la lectrice, du lecteur, avant même que ne se soit produit un premier contact avec l'écrit. Ce « déjà-là » (catégorisation d'objets, perception visuelle, capacités inférentielles, raisonnement), nécessaire au traitement de l'oral, est mobilisé dans les apprentissages implicites des régularités sur les plans orthographique, graphophonologique, morphologique, syntaxique, sémantique (Gombert, 2018).

Nous allons nous concentrer maintenant plus précisément sur le traitement de l'écrit, donc des processus de compréhension du discours écrit et de ses représentations, en accord avec les théories classiques du traitement du langage (Karimi et Ferreira, 2016). Au cours des 40 dernières années, un foisonnement de modèles ont vu le jour, la plupart s'inspirant de la proposition de Kintsch et van Dijk (1978) et van Dijk et Kintsch (1983), le *Contruction-Integration model*. Le propos n'est pas ici de faire un

panorama exhaustif des modèles de la compréhension (voir McNamara et Magliano (2009) pour une revue des modèles), mais de présenter les éléments convergents des approches théoriques sous-tendant la mesure de la compréhension.

4.2 Approches multicomponentielles de la compréhension

La compréhension de textes écrits constitue un processus cognitif complexe. La lectrice ou le lecteur doit procéder au traitement automatique et efficient de différents éléments linguistiques, ou niveaux de représentations, compris dans le discours : graphèmes, morphèmes, mots, syntaxe, proposition, message général (Gernsbacher et Kaschak, 2013; Graesser *et al.*, 1997). Pour comprendre, des connaissances sur la langue, mais aussi sur le monde, sont mobilisées (Gineste et Le Ny, 2002).

Le cadre dans lequel s'inscrivent les théories de la compréhension repose sur l'idée que le sens (d'une phrase, d'un texte, d'un discours), fondamentalement représentationnel, se construit grâce à un ensemble de processus (McNamara et Magliano, 2009). La construction du sens fait appel à des mécanismes d'échange entre inputs perceptuels et mémoire à long terme, situés en mémoire de travail (Graesser *et al.*, 1997; Le Ny, 2017).

Les approches cognitives multicomponentielles de la lecture cherchent à définir les composantes en jeu dans les performances en lecture (Hannon et Daneman, 2001, 2006). Généralement, les modèles de la compréhension en lecture font fonctionner

trois types de variables, toutes trois potentiellement sources de différences interindividuelles (Hannon, 2012; Hannon et Daneman, 2006) : des *représentations* (mots, propositions, image mentale, modèle de situation), des *processus* ou *mécanismes* de bas et de haut niveau (Lonigan, 2015) – notamment l'identification des mots et leur décomposition, mais aussi, et d'un point de vue connexionniste, l'activation, l'assemblage et la suppression (Gernsbacher, 1990, 1997) – et des *ressources cognitives* (en particulier la mémoire à long terme et la mémoire de travail (Baddeley, 1986; Just et Carpenter, 1992).

Alors que les ressources cognitives ne semblent pas particulièrement modifiables ou même conscientisables (Lonigan, 2015), certains processus et certaines représentations peuvent s'objectiver, notamment par l'entremise d'un travail métalinguistique ou d'un enseignement. Conséquemment, ce sont les représentations et les processus au sein de la compréhension en lecture qui retiendront notre attention dans la suite de cette revue de la littérature sur la compréhension en lecture.

4.3 Niveaux de représentation

Les théories et modèles de la compréhension en lecture partagent un intérêt marqué pour la nature des représentations qui jaillissent au fil de la lecture, lorsque le lecteur, la lectrice traite le discours (Karimi et Ferreira, 2016). Le discours porte en lui de multiples paliers de représentations ; phonèmes, morphèmes, syntagmes, propositions et message global sont traités rapidement et de manière coordonnée

(Graesser *et al.*, 1997). Les modèles théoriques de la compréhension du discours, et du discours textuel en particulier, ont été largement influencés par la notion de « niveaux » de représentation qui composent le processus de compréhension de textes (van Dijk et Kintsch, 1983), dans le sillage de la théorie des modèles mentaux, proposée par Johnson-Laird (1983). La notion de modèle mental est invariablement évoquée et mise en œuvre dans les modèles de la compréhension en général (Le Ny, 2017). Le construit « modèle mental » peut se référer à la représentation de tout type de connaissance systématisée, qu'elle soit organisée au sein d'une structure en 3D ou dans un format propositionnel avec une organisation syntaxique (Johnson-Laird, 2013).

van Dijk et Kintsch (1983) ont appliqué le concept de modèle mental à la compréhension du discours écrit, afin de représenter le discours et le monde qu'il dénote, ou la représentation qui est bâtie en référence à ce qui est lu. Avec le modèle de construction-intégration, Kintsch (1988, 1998) a poursuivi et étendu le modèle de Kintsch et van Dijk (1978) et van Dijk et Kintsch (1983). Le *Construction-Integration model* met en œuvre trois niveaux de représentations largement admis et mobilisés dans la littérature sur la compréhension de texte :

- La *structure de surface* du texte englobe les éléments lexicaux et syntaxiques explicitement mentionnés dans le texte qui sont reconnus.

- La *base propositionnelle* du texte correspond à la construction des inférences, à partir d'éléments explicites dans le texte, mais aussi à partir des connaissances que la lectrice ou le lecteur possède en mémoire.
- Le *modèle de situation* du texte fait référence au modèle mental spécifique qui est créé dans un mouvement d'interaction entre les informations implicites fournies par le texte et les connaissances stockées en mémoire à long terme que la lectrice ou le lecteur a sur le monde. Il s'agit du « micromonde » généré par la lecture et dénoté par le texte (Graesser *et al.*, 1997).

Le texte en lui-même est finalement assez peu explicite (Farrall, 2012). La lecture d'un texte et la construction de sa compréhension se situent fondamentalement dans l'implicite. C'est la présentation de toutes les informations et des propositions nécessaires à une interprétation cohérente du texte qui est une entité théorique (van Dijk et Kintsch, 1983). Les informations qui manquent et qui permettent de rendre le discours cohérent sont fournies par le lecteur, qui s'appuie sur ses représentations cognitives ou sémantiques du contexte et des faits, emmagasinées en mémoire à long terme (Gineste et Le Ny, 2002; Snow *et al.*, 1998; van Dijk et Kintsch, 1983).

4.3.1 Structure de surface (*surface code*)

La structure de surface correspond à une représentation linguistique immédiate de l'input textuel (phrases, mots, syntaxe). Les mots du texte sont des nœuds reliés par des relations syntaxiques (McNamara et Magliano, 2009). Le lecteur ou la lectrice ne

retient habituellement que les éléments les plus récemment lus. Ces éléments ne sont retenus que si leur réinvestissement assure la construction du sens global du texte (Graesser *et al.*, 1997). Ce niveau n'est pas fondamental dans la construction de la compréhension, et a même été abandonné dans le modèle proposé par Kintsch (1998).

4.3.2 Base propositionnelle du texte (*textbase*)

La base propositionnelle du texte représente les informations sous-jacentes aux informations explicitement contenues dans le texte. Il s'agit d'une structure essentiellement sémantique, dans laquelle les phrases et les mots du texte ont disparu pour laisser place aux propositions issues du texte et de quelques inférences assurant la cohésion textuelle (Graesser *et al.*, 1997; Kintsch, 1998).

Empruntées à la logique, les propositions sont généralement considérées comme des unités sémantiques cognitives fondamentales sur lesquelles il est possible de réfléchir, raisonner dans les champs de la psychologie cognitive de la philosophie et de l'intelligence artificielle (Farrall, 2012; Gineste et Le Ny, 2002). Elles constituent par ailleurs la notion la plus proche de celle d'évènement, si on veut la rapprocher du genre narratif, ou d'idée, d'un point de vue plus général. Les propositions sont constituées d'au moins un prédicat (par exemple, le verbe principal, les connecteurs, les adjectifs) et de ses arguments (noms et propositions enchâssées) (Gineste et Le Ny, 2002; Goldman et Wolfe, 2001; Graesser *et al.*, 1997). Les arguments sont ici des

choses, et le prédicat exprime une propriété qu'ont ces choses ou le lien qui les unit (Ehrlich et Tardieu, 1993). Leur structure canonique est la suivante : prédicat (argument) (argument) (Gineste et Le Ny, 2002; McNamara et Magliano, 2009). Elles constituent des « unités prédicatives », des « unités de pensée » (Ehrlich et Tardieu, 1993, p. 52), qui correspondent à des représentations d'états, d'évènements ou d'actions possédant une valeur de vérité. Cette caractéristique leur confère le pouvoir d'être traitées comme des arguments dans les formes de raisonnement, notamment déductif (Johnson-Laird, 2013).

La base propositionnelle de texte se compose d'une *microstructure* et d'une *macrostructure*, toutes les deux fonctionnant avec les propositions comme unités de base (Farrall, 2012; Kintsch, 1998; Kintsch et van Dijk, 1978). La microstructure fait référence à la structure propositionnelle locale du texte, obtenue à partir des informations des phrases, et telle que complétée par l'interaction avec les informations contenues en mémoire à long terme. La macrostructure dérive de la microstructure. Elle fait référence à l'organisation globale du texte, et correspond à une structure hiérarchiquement ordonnée des propositions représentant l'organisation du texte dans son entièreté (Kintsch, 1998). La microstructure et la macrostructure sont dans une relation d'orthogonalité avec les notions de base propositionnelle textuelle et de modèle de situation (Kintsch, 1998, p. 50). Ainsi, la base propositionnelle du texte, avec ses micro et macrostructures, est obtenue à partir d'une analyse sémantique d'un texte et de sa structure rhétorique, telle que l'auteur·e du texte l'a voulue, mais

par l'interaction avec la mémoire à long terme de la lectrice ou du lecteur. La macrostructure n'est pas susceptible d'être fortement éloignée de la microstructure, si le texte présente des idées organisées sériellement (McNamara et Magliano, 2009).

Afin de consolider par l'exemplification les notions préalablement exposées, considérons les représentations mises en place pour arriver à une compréhension du passage suivant (en 1), tiré de *Star Wars Épisode IV : Un nouvel Espoir* (1977) :

(1)

> ***Il y a bien longtemps, dans une galaxie lointaine, très lointaine...***
> *a) C'est une époque de guerre civile.*
> *b) À bord de vaisseaux spatiaux opérant à partir d'une base cachée, les Rebelles ont emporté leur première victoire sur le maléfique Empire galactique.*
> *c) Au cours de la bataille, des espions rebelles ont réussi à dérober les plans secrets de l'arme absolue de l'Empire : l'Étoile de la Mort, une station spatiale blindée dotée d'un équipement assez puissant pour annihiler une planète tout entière.*
> *d) Poursuivie par des sbires sinistres de l'Empire, la princesse Leia file vers sa base dans son vaisseau cosmique, porteuse des plans volés à l'ennemi qui pourront sauver son peuple et restaurer la liberté dans la galaxie...*

Dans l'exemple en (1), la représentation propositionnelle de la phrase a), qui est un peu différente de celles qui sont données en exemple normalement, à cause du verbe copule, est la suivante :

P1 : civile (OBJET = guerre) 'La guerre est civile'

P2 : P1(guerre) (OBJET = époque) 'Il s'agit d'une époque de guerre'

Dans cette phrase simple en (1a), et d'un point de vue cognitif, le signifiant 'époque' est complété par la proposition 'guerre civile'. Le lien de cohésion entre la phrase 1b) et la phrase 1c) est renforcé par le chevauchement d'arguments, dans ce cas-ci, « rebelle » et « empire ». L'inférence est locale. Pour ce qui est du rapprochement entre les phrases 1a) et 1b), il n'est pas assuré par la présence d'un argument commun au sein des deux propositions. Ce lien est plutôt assuré par l'établissement d'inférences basées sur des connaissances du monde (*knowledge-based inferences*) (McNamara et Magliano, 2009), un niveau représentationnel correspondant au modèle de situation du texte.

Il est à noter cependant que l'élaboration d'un modèle de situation pourrait ne pas dépendre de manière critique de la présence ou de l'absence de chevauchement d'arguments (Zwaan, Magliano et Graesser, 1995), ce qui met à mal l'importance de la base propositionnelle du texte comme unité distincte de représentation textuelle, une idée mise de l'avant en particulier par Kintsch et van Dijk (1978). Il semble que les lectrices et lecteurs sont en mesure de piloter simultanément plusieurs dimensions globales du modèle de situation, notamment la temporalité et la causalité, sans égard à la cohésion locale (Zwaan, Magliano et Graesser, 1995).

4.3.3 Modèle de situation

La compréhension d'un texte sous-tend certes la construction d'une représentation mentale des mots, des phrases et des arguments, d'ordre linguistique, mais aussi de

la représentation de la situation portée par ses éléments linguistiques, qui est appelée modèle de situation (Zwaan, Magliano et Graesser, 1995). Cette notion est constamment intégrée aux recherches sur le discours (Kintsch, 1998; Kintsch et van Dijk, 1978; van Dijk et Kintsch, 1983). Il s'agit du « micromonde » dénoté par le texte (Graesser *et al.*, 1997, p. 167; Zwaan, 2001), et constitue le niveau le plus élaboré de la compréhension. Le modèle de situation est un type de modèle mental spécifique à la situation et aux évènements décrits dans le texte (Zwaan, 2001). Alors que les modèles mentaux correspondent à des représentations de situations réelles ou hypothétiques, et souvent plus générales, les modèles de situation représentent des évènements, dans leurs dimensions temporelles, spatiales, causales et motivationnelles (McNamara et Magliano, 2009; Zwaan, 2001). Selon le modèle de la compréhension en lecture de Gernsbacher (1990), le *Structure-Building Framework*, l'objectif de la compréhension des informations linguistiques et non linguistiques est de construire une représentation mentale cohérente, ou « structure ». Le concept de « structure » dans le modèle rappelle celui du modèle mental de Johnson-Laird (1983) et s'en inspire. Il a été spécifiquement développé pour la compréhension de textes narratifs.

Le concept de modèle de situation donne corps à la notion de compréhension profonde (McNamara et Magliano, 2009), en reflétant une compréhension qui va au-delà de ce qui est dénoté explicitement par le texte ou dans la structure propositionnelle en elle-même. Il s'agit d'une représentation mentale, ou de l'état en

cours de la structure créée (Gernsbacher, 1990), cohérente de la situation dénotée par le récit (Fayol, 2003; Gernsbacher, 1990, 1991; Kintsch et van Dijk, 1978; Oakhill et Garnham, 1988; van Dijk et Kintsch, 1983). Le modèle de situation représente véritablement la situation décrite par le texte, et non le texte lui-même. Il est réputé exempt de représentations proprement linguistiques (Zwaan, 2001).

La compréhension du discours réside fondamentalement dans la recherche d'implicite ; la présentation de toutes les informations et des propositions nécessaires à une interprétation cohérente du texte n'est effectivement qu'une entité théorique (van Dijk et Kintsch, 1983). Le modèle de situation constitue dans ce cadre une construction du discours à partir d'informations qui ne sont pas directement connectées entre elles, en interaction constante avec les affects de la lectrice ou du lecteur. La construction du modèle suppose la création et le maintien en mémoire de schèmes de pensée chez la personne qui lit. Celle-ci s'attend à ce que ce qu'elle lit soit cohérent : elle fournit les informations qui manquent et qui permettent de rendre le discours cohérent, en s'appuyant sur sa connaissance du contexte et des faits (van Dijk et Kintsch, 1983).

La compréhension linguistique est intrinsèquement difficile à modéliser, notamment computationnellement, à cause de son caractère complexe, mais aussi parce qu'elle implique l'interaction de l'être humain avec son environnement (Zwaan, 1999) et que l'environnement inclut les perceptions très personnelles comme les émotions et

l'histoire individuelle (Marinier et Laird, 2006). Il est difficile de penser « compréhension » sans penser « différences individuelles », tant cette habileté implique le tissage de liens avec l'histoire personnelle de la lectrice ou du lecteur.

Nous pouvons élaborer une liste d'inférences possibles, parmi d'autres, tout aussi valides, qui peuvent contribuer à la construction du modèle mental à partir de l'exemple en (1) :

- 'La princesse est en danger.'
- 'La princesse, malgré son statut royal, fait partie du camp des Rebelles.'
- 'L'Empire est en position de force.'
- 'Entre les deux camps, le mal est représenté par l'Empire, et le bien, par les Rebelles.'

Les inférences participent du modèle mental autant qu'ils sont fonction des modèles mentaux que la lectrice ou le lecteur a préalablement construits (McNamara et Magliano, 2009).

4.4 Processus

Nous avons vu que la compréhension de l'écrit implique des opérations cognitives à tous les niveaux de la représentation linguistique. Elle engage virtuellement tous les

processus cognitifs, comme la perception, la mémoire, le raisonnement et la résolution de problèmes (Graesser *et al.*, 1997).

On distingue habituellement les processus de bas niveau des processus de haut niveau. Les processus de bas niveau impliquent la manipulation d'unités plus proches du matériau écrit (mots, syntaxe). Les processus de haut niveau, susceptibles de prendre une plus grande part de variance dans les résultats en compréhension (Hannon et Daneman, 2001), sont réputés être *domain-general*, c'est-à-dire que les mécanismes sous-jacents de la construction du modèle ou de la structure sont considérés comme des processus et mécanismes cognitifs généraux (une position particulièrement mise de l'avant dans le *Structure-Building Framework* de Gernsbacher, 1990). Le modèle CCR, pour *Cognitive Components-Resource Model of Reading* (Hannon, 2012, 2016; Hannon et Daneman, 2001), met particulièrement l'accent sur le fait que les processus cognitifs de bas niveau et de haut niveau sont des constructions distinctes.

4.4.1 Approches ascendante, descendante et interactive de l'acte de lire

Trois types de modèles impliquant les processus de bas et de haut niveaux ont vu le jour pour rendre compte des mécanismes cognitifs relatifs à l'acte de lire : ascendant, descendant et interactif (de Weck et Marro, 2010).

Les approches *bottom-up* en lecture mettent de l'avant l'hypothèse que le traitement des représentations se fait de manière ascendante, soit à partir des informations de bas niveau jusqu'à celles de haut niveau (Ferrand et Spinelli, 2005). Les représentations phonologiques et graphiques sont identifiées, donnant accès aux représentations lexicales. Celles-ci sont associées à des représentations sémantiques. Ces informations sont réputées être stockées en mémoire jusqu'à ce qu'elles soient associées à des unités linguistiques plus macro, organisées en fonction de règles : groupe syntaxique, phrase, texte (de Weck et Marro, 2010). Les méthodes pédagogiques alphabétiques et syllabiques (Borel-Maisonny, 1956) sont issues de cette approche.

Dans les approches descendantes, dites *top-down*, il est entendu que la lectrice ou le lecteur active ses connaissances du monde et des discours d'abord, afin de comprendre un texte (Ferrand et Spinelli, 2005). Les modèles *top-down* suggèrent que la construction du sens se fait par un mécanisme de sélection efficace d'indices et de prédictions sémantiques et syntaxiques. Le recours aux mécanismes de bas niveau pendant la lecture n'est effectué que pour vérifier la validité de ces prédictions (de Weck et Marro, 2010). Cette conception de l'acte de lire se matérialise pédagogiquement par l'avènement des méthodes globales (Goodman, 1989).

Ces conceptions opposées de la lecture, et leur échec à décrire et expliquer les performances de l'ensemble des lectrices et lecteurs (de Weck et Marro, 2010), ont

donné lieu à l'avènement des approches interactives (Goigoux, 1998), qui donnent une large place à deux composantes de la lecture : le décodage graphophonologique, réputé spécifique à l'écrit, et la compréhension, vue comme partageant les mêmes mécanismes à l'oral qu'à l'écrit (Gough et Tunmer, 1986).

Le modèle simple de la lecture (*Simple view of reading*) (Gough et Tunmer, 1986; Hoover et Gough, 1990) est un modèle bidimensionnel du développement de la lecture constamment mis de l'avant dans les travaux s'intéressant à la litéracie et au développement des habiletés en lecture (Bianco *et al.*, 2012; Bourdages, 2014; Hoover et Tunmer, 1993; Joshi et Aaron, 2000; Kendeou *et al.*, 2016; Park et Uno, 2015; Perfetti *et al.*, 2005; Spencer, M. et Wagner, 2018; Tunmer et Chapman, 2012) qui donne corps au postulat selon lequel la compréhension du code écrit est le produit du décodage et de la compréhension linguistique orale. Cette position donne lieu à la fameuse formule L = D x C, mis pour Lecture = Décodage x Compréhension linguistique (*Reading* = *Decoding* x *Comprehension*, R = D x C), où chaque variable est comprise entre 0 et 1 (Gough et Tunmer, 1986). Le modèle sous-tend que les deux dimensions contributives sont nécessaires et qu'elles ne peuvent fonctionner indépendamment l'une de l'autre.

Régulièrement remis en question (cf. notamment Foorman *et al.* (2015), Hannon (2012) et Compton *et al.* (2013) à ce sujet), le modèle fait la prédiction, vérifiée (Aarnoutse *et al.*, 2001), que les corrélations entre le décodage et la compréhension

de la lecture diminueront avec le développement, alors que les corrélations entre la compréhension linguistique et la compréhension de la lecture augmenteront (Hoover et Gough, 1990).

Le modèle permet aussi de tenir compte du chevauchement entre la langue écrite et la langue orale dans leur structure de base. Dans les langues à système d'écriture alphabétique, comme le français, l'importance relative du décodage et de la compréhension orale est fonction de la transparence de l'orthographe. La formule du modèle permet de faire la prédiction selon laquelle les langues à orthographe transparente induiront une importance plus forte de la compréhension orale (C) sur les performances en lecture (L) (Lonigan, 2015).

La formule du modèle permet également de schématiser les différences interindividuelles dans les performances en lecture. Il existe en effet une littérature considérable qui fait état de l'existence de groupes d'enfants – environ 10% (Farrall, 2012) – qui ont une faible compréhension en lecture malgré des habiletés en décodage, en fluidité, en lecture de mots isolés, en conscience phonologique et en vocabulaire adaptées à leur âge (voir Cain et Oakhill (2007) pour une revue de la littérature au sujet de cette population). Le modèle prédit que si les mécanismes *bottom-up* sont faibles et accomplis avec effort, la compréhension sera entravée, puisque les mots sont mal identifiés et parce que généralement moins de ressources cognitives pourront être consacrées au traitement du sens et à des opérations

cognitives plus complexes (Perfetti, 1985). Lorsque les mécanismes *top-down* font défaut, et ce, malgré des habiletés intactes en décodage, la compréhension du texte n'est pas effective, et l'accès au sens des mots n'est pas atteint, les relations logiques et structurelles entre eux et entre les propositions sont mal ou peu établies (Nation et Snowling, 2000).

4.4.2 Processus de bas niveau

Les modèles de la compréhension du discours considèrent ces processus de bas niveau, spécifiques à la situation d'un texte écrit, comme essentiels : ils constituent des prérequis à la mise en place des processus de haut niveau, sans en être nécessairement précurseurs (en temps réel, et non du point de vue développemental ou de sérialité des processus en jeu dans la lecture). Les processus de bas niveau évoqués sont considérés comme déjà automatisés et bien en place au moment où s'opèrent les processus de haut niveau inclus dans les modèles de la compréhension (McNamara et Magliano, 2009, p. 302).

4.4.2.1 Décodage

Au niveau le plus élémentaire, la lecture exige de la lectrice ou du lecteur qu'elle ou il opère un décodage graphophonologique des symboles graphiques, reconnaisse les mots et procède au *parsing* ou analyse syntaxique des phrases (Gernsbacher et

Kaschak, 2013). Le décodage se définit comme la capacité automatisée à transformer un input graphique en un code phonémique (Perfetti, 1985).

4.4.2.2 Accès au lexique

La réussite au processus de décodage, ou d'appariement des représentations phonémique et graphémique, permet l'automatisation de l'accès au lexique et de ses représentations orthographiques et lexicales, et leur stockage en mémoire (Ehri, 2005). Il préside donc à l'identification de suites de signes graphiques lus peu fréquemment ou pour la première fois (mots nouveaux, mots rares). Il s'agit d'un mécanisme de perception (traitement visuel) relativement peu couteux cognitivement, impliquant les informations fovéales et parafovéales et de codage orthographique prélexical (de Weck et Marro, 2010).

Les travaux de psychologie cognitive sur l'implication du décodage dans l'accès au lexique s'appuient principalement sur des tâches de décision lexicale et de catégorisation sémantique, qui impliquent des temps de réponse (Ferrand et Spinelli, 2005; Rastle, 2009). Parmi eux, le *Dual Route Cascaded Model* (modèle DRC) proposé par Coltheart *et al.* (2001), modèle d'inspiration symbolique[52], permet d'établir

[52] Les architectures de type symbolique permettent de bien circonscrire l'habileté en décodage. En effet, conceptuellement, le décodage est une sorte de calcul très simple de relations plus ou moins 1 : 1, où les règles de correspondances sont des règles d'interprétation. La fonction d'interprétation fonctionne par la computation des symboles écrits, desquels on hérite des symboles phonologiques.
Cependant, des modèles connexionnistes comme ceux de Seidenberg et McClelland (1989) ou de Plaut *et al.* (1996) ont permis de faire la simulation de quantité de données sur le

un lien entre les représentations phonologiques et sémantiques et la reconnaissance des mots écrits, et a, par ailleurs, permis la simulation informatique, dans une perspective de validation théorique (Coltheart, 2005; Rastle, 2009). Selon ce modèle, la lectrice ou le lecteur accède à la représentation du mot via une voie directe ou indirecte : la première passe par un traitement orthographique, global, du mot, l'autre est indirecte et opère par assemblage (à un graphème son phonème) (Coltheart, 2005; Coltheart *et al.*, 2001). Ce modèle fait donc intervenir explicitement le décodage au cours de la lecture de mots n'ayant pas encore d'entrée orthographique dans le lexique mental : la lectrice ou le lecteur habile peut utiliser les deux routes, alors que les débutant·es sont contraint·es d'utiliser la voie indirecte (Treimen et Kessler, 2009). Des connexions facilitatrices et inhibitrices sont postulées pour expliquer comment les modules communiquent entre eux dans le modèle.

Ces modèles ne représentent qu'une partie des processus impliqués dans l'acte de lire, soit ceux qui concernent les habiletés à automatiser pour identifier habilement et efficacement les mots écrits ; ceux qui concernent spécifiquement la compréhension seront traités dans la section suivante.

développement de la lecture de mots et de rendre compte de plusieurs effets observés en lecture. Le modèle de reconnaissance des mots de Seidenberg et McClelland (1989) implique également le calcul des codes orthographique, phonologique et sémantique, ainsi que l'encodage d'informations sur la fréquence et la cohérence des correspondances entre l'orthographe et la phonologie.

4.4.3 Processus de haut niveau

Les processus de haut niveau sont ceux qui sont en relation avec l'existence et la construction d'un modèle mental[53] (Johnson-Laird, 1983), qui, pour rappel, dénote la représentation mentale que la lectrice ou le lecteur se construit en référence à ce qu'elle ou il lit.

> Vous construisez des modèles du monde à partir des processus perceptuels, mais vous pouvez également les construire à partir de descriptions du monde, qui vous permettent d'en faire l'expérience par procuration. Un bon écrivain ou conteur a le pouvoir d'initier un processus similaire à celui qui se produit lorsque vous percevez ou imaginez des évènements. [notre traduction de Johnson-Laird (2013, p. 3)] [54]

Le modèle mental est conçu dans ce cadre comme le produit psychologique du raisonnement humain, et sa construction est fonction de l'échange entre processus perceptifs, connaissances antérieures stockées en mémoire à long terme et capacités de la mémoire de travail.

---

[53] Un modèle mental n'est pas un modèle au sens où nous l'entendons le plus souvent dans le cadre de cette thèse, c'est-à-dire dans un sens proche de « théorie » ; il s'agit d'une structure psychologique interne, comme pourrait l'être une image mentale.

[54] « *You construct models of the world from perception, but you can also construct them from descriptions of the world, which enable you to experience it by proxy. A good writer or storyteller has the power to initiate a process similar to the one that occurs when you perceive or imagine events.* »

4.4.3.1 Construction, appariement et modification des modèles mentaux

Un des modèles les plus productifs et intéressants mettant en œuvre les processus mentaux oeuvrant à la construction des modèles mentaux est le modèle de Gernsbacher (Gernsbacher, 1990, 1991, 1995, 1997; Gernsbacher *et al.*, 1990). Ce modèle met l'accent sur les processus mis en place pour la construction des *structures* pendant la compréhension. La construction de ces structures repose sur des mécanismes cognitifs généraux, qui fonctionnent sans égard au type d'informations traitées (McNamara et Magliano, 2009). Le modèle garantit une compréhension affranchie de la modalité qui la caractérise, et donc de mécanismes de plus bas niveau. Trois processus généraux assurent la création des structures : (1) la construction itérative des fondements de la structure (*laying a foundation*), (2) l'appariement des nouvelles informations au sein de la structure en formation (*mapping*) (répétition de concepts, chaine anaphorique et autres indices syntaxiques), et (3) la modification de la structure lorsque la cohérence est brisée, c'est-à-dire lorsqu'une information nouvelle ne peut être appariée à la structure pour cause d'incongruence, ou manque de cohérence[55], avec la structure préétablie, ou parce que la nouvelle information correspond au début d'une nouvelle idée (Gernsbacher, 1997).

[55] La *cohérence* se manifeste par les connexions établies entre les propositions et par l'activation des connaissances antérieures et leur incorporation à de nouvelles représentations mentales. Elle est relative à la lectrice ou au lecteur, puisqu'il est fortement influencé par le modèle de situation. La *cohésion*, quant à elle, se réfère au texte : elle consiste en le chevauchement des arguments au sein de celui-ci (Mc Namara et Magliano, 2009).

4.4.3.2 Activation et assemblage

Ces processus à la base de l'affinement des structures ne sont pas sans rappeler ceux du modèle construction-intégration (Kintsch, 1998), dans lequel les connaissances sont activées dans la phase de construction, puis les représentations activées subissent un processus itératif de satisfaction de contraintes durant la phase d'intégration (McNamara et Magliano, 2009).

Ces processus généraux sont fonction du niveau d'activation des nœuds mémoriels (*memory nodes*). Le niveau d'activation est tour à tour augmenté et réduit par deux mécanismes minimaux : le renforcement (*enhancement*) et la suppression (*suppression*) (Gernsbacher, 1997). Ces mécanismes assurent la construction et le maintien en mémoire des structures et sous-structures, ou représentations mentales. La compréhension est ainsi conçue comme dynamique, active : les informations fournissent les éléments nécessaires à la construction d'une structure interprétative inédite. Celle-ci est susceptible de changer et de s'affiner au fil de la lecture, étant confrontée à des informations nouvelles qui la valident ou l'invalident.

Les lectrices et lecteurs sont conscient·es du produit de la compréhension, et sont même en mesure de diriger leurs ressources attentionnelles vers les informations textuelles pertinentes, en fonction des objectifs visés (McNamara et Magliano, 2009). Les modèles de la compréhension incluent aussi des mécanismes non conscients et automatiques mis en œuvre durant la lecture. Ces processus ne sont pas disponibles

pour la verbalisation, ne sont à priori pas conscientisables. Du point de vue des architectures connexionnistes, dont font partie la grande majorité des modèles (McNamara et Magliano, 2009), l'activation des nœuds (*nodes*) représentationnels est un prélude aux processus subséquents assurant la compréhension (Gineste et Le Ny, 2002).

Il y a ainsi activation de significations préexistantes (lexicales notamment), et assemblage (règles morphologiques, syntaxiques, sémantiques) (Le Ny, 2017). L'activation d'un concept active ou abaisse le seuil d'activation de concepts reliés. La propagation de l'activation est appliquée aux concepts disponibles en mémoire. Ce processus est réputé modifier l'activation des concepts en fonction de leur connectivité et de leur poids initial dans la représentation (McNamara et Magliano, 2009).

4.4.3.3 Raisonnement déductif et inférences

Il est impératif de décrire les mécanismes permettant le raisonnement logique puisque celui-ci est une partie intégrante de l'esprit humain (Johnson-Laird, 2013). La psychologie cognitive jette un regard nouveau – et différent, par rapport à celui de la philosophie classique – sur la logique, le raisonnement et les inférences, dans la mesure où il est généralement admis que l'esprit humain fonctionne probablement en partie grâce à un système de déduction (Johnson-Laird, 2010). Ainsi, le raisonnement est fonction de la mémoire de travail, de l'attention et d'autres fonctions exécutives.

Les inférences, soit les évènements et états qu'il est nécessaire d'établir afin de garantir une cohérence plus générale du texte, et sur le raisonnement, comme calcul des propositions (Johnson-Laird et Khemlani, 2014), sont des composantes de la construction du modèle mental (Ehrlich et Tardieu, 1993; Johnson-Laird, 2013). Certains modèles, notamment le *Causal Network model* (Trabasso *et al.*, 1989), suggèrent même que la construction du modèle mental, comme représentation cohérente du texte, est principalement fondée sur le raisonnement causal. Les connaissances sont, dans ce cadre, des éléments permettant d'en arriver à des raisonnements basés sur l'expérience ou sur une connaissance préalable du monde en général. La mémoire de travail est conçue comme le lieu où l'intégration des connaissances et les inférences ont lieu (Cain *et al.*, 2004).

Le raisonnement est traditionnellement considéré comme proche de la notion d'inférence (Campion et Rossi, 1999), dans le sens où il est généralement admis que la fonction du raisonnement est de produire des inférences correctes à partir des inputs sensoriels et de l'environnement (Ehrlich et Tardieu, 1993). Le raisonnement est un élément extrêmement important dans l'expérience humaine[56]. La psychologie

[56] Selon la théorie du langage de la pensée (LOT, mis pour *Language of Thought* en anglais), dont le principal porte-étendard est Jerry Fodor (1981), les représentations mentales possèdent une structure analogue à la structure linguistique, notamment une syntaxe et une sémantique combinatoire. Les propositions (*p,q,r*) possèdent des valeurs de vérité et constituent des représentations primitives qui se complexifient par le calcul des propositions, comme dans l'implication de type $p \Rightarrow q$, où *p* implique *q,* est égal, sur le plan logique, à $(\neg p) \vee q$ (non-*p* ou *q*). L'idée que le raisonnement humain peut être expliqué par la logique traditionnelle est contestée, notamment parce que la déduction, par exemple, peut mener à des conclusions farfelues, et traite d'arguments conditionnels sans tenir compte des probabilités (Johnson-Laird *et al.*, 2015).

du raisonnement étudie l'habileté humaine à faire des inférences, de type déductif, où les prémisses fournissent des raisons permettant d'accepter la conclusion, ou inductif, où les prémisses ne fournissent pas de raisons permettant d'accepter la conclusion (celui-ci n'est pas basé sur un raisonnement probabiliste). Ces deux types d'inférences partagent certains mécanismes, mais pas tous[57]. Le *modus ponens* en est un exemple générique issu de la logique classique : il s'agit d'une opération déductive symbolique purement computationnelle. Il correspond à une inférence logique qui se traduit schématiquement comme suit : « si *P* implique *Q*, et que *P* est vrai, alors *Q* est vrai ».

Le raisonnement logique[58] est assuré par de multiples systèmes distribués (systèmes du langage, des sens, des mécanismes moteurs, des émotions, etc.), réagissant aux stimuli environnementaux dans une relation dynamique, et qui tient en des buts, des croyances et des actions.

Les inférences et le raisonnement présentent l'avantage d'être aisément conscientisés, objectivés, enseignés. Leur réalité psychologique, qui consiste globalement en la construction d'unités de sens plus larges – des idées – à partir des

[57] Ces processus cognitifs pourraient même avoir une réalité neurologique. En effet, le cortex préfrontal semble mobilisé pendant le raisonnement, notamment dans le cas de syllogismes, ou *modus ponens*, et des inférences probabilistes (Parsons et Osherson, 2001). Le raisonnement est fonction des systèmes linguistique et spatial. Il semble que les inférences simples aient pour origine le système permettant le langage et que les inférences plus complexes soient prises en charge par le système visuospatial (Goel *et al.,* 2000).

[58] On entend ici la logique au sens d'un langage formel permettant de faire des inférences valides.

unités de sens contenues dans le texte – les propositions –, constitue un élément de convergence pour les modèles de la compréhension[59] (McNamara et Magliano, 2009).

4.5 Mesure de la compréhension en lecture

L'utilisation d'outils de mesure fiables de la compréhension en lecture constitue un enjeu pour les chercheur·es et praticien·nes de disciplines diverses, notamment les orthophonistes, enseignant·es et psycholinguistes. En effet, la compréhension en lecture, construit reconnu comme gage de réussite scolaire, personnelle et professionnelle, revêt un intérêt majeur pour qui s'intéresse à l'apprentissage et à l'enseignement (Berger, M. J. et Desrochers, 2011; OCDE, 2001, 2014; Vaughn *et al.*, 2015). Les recherches sur la compréhension en lecture ont tôt été happées par la tension entre deux objectifs, pas antinomiques, mais distincts et susceptibles de s'entrechoquer : d'une part, la mesure, et d'autre part, la théorie (Hannon et Daneman, 2001; Rupp, 2012).

4.5.1 Perspectives diagnostique et explicative de la compréhension

Pour une grande part, c'est dans une perspective diagnostique (Bourdages, 2014) que le construit a été investi par les chercheur·es. Cet intérêt pour une mesure de la

[59] Ces modèles, pour l'écrasante majorité connexionnistes (McNamara et Magliano, 2009), échouent pourtant à intégrer les comportements délibérés et conscients, donc métacognitifs, propres à ceux que sont en mesure de mettre en place les lecteur·ices.

compréhension qui soit prédictive des troubles en lecture a fait naitre de nombreuses batteries de tests standardisés, parmi lesquelles nous citerons les plus connues en anglais : *Test of silent reading efficiency and comprehension* (TOSREC) (Wagner *et al.*, 2010), *Nelson-Denny Reading Test* (Brown, J. I. *et al.*, 1993), *Gates-MacGinitie Reading Test* (MacGinitie et MacGinitie, 1989; MacGinitie *et al.*, 2000), *Passage comprehension subtest* du Woodcock-Johnson III (WJ-III) (Woodcock *et al.*, 2001), *Woodcock reading mastery tests—Revised* (WRMT-R) (Woodcock, 1987) et en français : Batterie Analytique du Langage Écrit (BALE) (Jacquier-Roux *et al.*, 2010), Batterie de Lecture et de Compréhension de Récit (BLCR) (Alexandre-Beaumont *et al.*, s.d.) et Lecture de mots et compréhension, revisée (LMC-R) (Khomsi, 1999).

Nous avons préalablement montré que les différents modèles de la compréhension, notamment le modèle construction-intégration (*CI*) (van Dijk et Kintsch, 1983), le *Structure-Building Framework* (Gernsbacher, 1990), et le *Cognitive Components-Resource model* (Hannon et Daneman, 2001), mettent tous en œuvre des niveaux de représentation et des processus. Ces niveaux de représentations, qui peuvent aussi être conçus comme des produits de la compréhension (McNamara et Magliano, 2009) sont inégalement mis en œuvre dans les études empiriques (Farrall, 2012). L'un des défis persistants dans la mesure de la compréhension a été de concevoir des tâches expérimentales qui permettent de considérer isolément chacun des différents niveaux de représentation (Graesser *et al.*, 1997).

Dans une perspective explicative, un large éventail de design expérimentaux ont été mis en place avec pour objectif de faire la lumière sur ces mécanismes mobilisés lors de la compréhension. Si les différentes approches théoriques sur la compréhension en lecture incluent toutes des niveaux de représentation – la structure de surface du texte, la base propositionnelle du texte et le modèle de situation, ou modèle mental (Gernsbacher, 1997; Kintsch, 1998; Kintsch et van Dijk, 1978) – les épreuves permettant effectivement de juger des performances en compréhension en lecture échouent à mesurer l'habileté à construire un véritable modèle mental (Farrall, 2012; Hannon, 2012). Hannon et Daneman (2006, p. 130) qualifient même d'« athéoriques » la plupart des batteries d'épreuves censées mesurer la compréhension. Crucialement, enfin, notons que les études portant sur les liens entre les habiletés métalinguistiques et la compréhension en lecture n'utilisent le plus souvent qu'une épreuve de compréhension dans leur modèle prédictif, et celle-ci est souvent constituée de questions de compréhension simples, et avec réponses à choix, sur de courts passages de textes (*passage comprehension*) – voir notamment le *Nelson-Denny Reading Test* (Brown, J. I. *et al.*, 1993). La mesure du modèle de situation est plus difficile (Bourdages, 2014), ce qui explique la relative rareté des tentatives dans la littérature scientifique (Gernsbacher, 1991; Graesser *et al.*, 1997; Hannon, 2012; Zwaan, Langston et Graesser, 1995).

4.5.2 Épreuves de compréhension dans les études sur la CM et la CS

Les publications psycholinguistiques qui tentent de lier expérimentalement des habiletés spécialisées avec des mesures de performances langagières plus générales ont le vent en poupe. La présente thèse en est d'ailleurs un exemple flagrant, puisque nous tentons de mettre en rapport les habiletés en conscience morphologique et en conscience syntaxique avec la qualité de la compréhension de la lecture. Nous avons d'ailleurs fait état, dans les précédents chapitres, de notre revue de la littérature récente sur ces liens. Nous pouvons faire deux constats à la lumière de cette revue méthodologique. D'abord, la part que prend la compréhension en lecture dans la mesure de ce que les auteur·es appellent souvent indistinctement « lecture » (ou « *reading* ») est soit absente, soit prise en compte dans une mesure de la lecture qui inclut d'autres construits réputés constituer ce construit, comme le vocabulaire ou le décodage (Farrall, 2012). D'autre part, il existe un réel décalage entre les outils de mesure élaborés pour juger du niveau de compréhension ou des performances en compréhension et les modèles théoriques anciens et attestés (van Dijk et Kintsch, 1983) ou relativement récents (Hannon et Daneman, 2001).

Le tableau 4.1 rend compte des tests de compréhension les plus couramment utilisés dans les travaux sur les liens entre habiletés métalinguistiques et lecture. Notons que seules les épreuves censées mesurer la compréhension, et non le décodage, ont été recensées. Cette clarification est importante, dans la mesure où il n'est pas rare de

rencontrer des mesures de la compréhension en lecture qui soient le produit de mesures autres que la compréhension.

Tableau 4.1 Vue d'ensemble des tests de compréhension les plus couramment utilisés dans les travaux sur les liens entre habiletés métalinguistiques et lecture

Tests de compréhension	Forme du test
Test of Silent Reading Efficiency and Comprehension – TOSREC (Wagner *et al.*, 2010)	Lecture silencieuse de texte, déterminer si les phrases sont vraies ou fausses (*yes/no*)
Sous-test *Reading Comprehension* du *Comprehensive Testing Program* – CTP, Comprehension Testing Program III, *Educational Records Bureau* (Carlisle, 2000)	Lecture de courts passages et question de compréhension à choix
Nelson-Denny Reading Test – Form G, part 3 - 7 (Brown, J. I. *et al*., 1993)	Passages, lecture silencieuse, 38 questions avec cinq réponses à choix
Sous-test de compréhension en lecture (niveau 3) de Gates-MacGinitie – 4e édition (MacGinitie *et al*., 2000)	Lecture silencieuse de 11 passages de longueurs variables 48 questions à choix multiples
Gates–MacGinitie Reading Comprehension Test (MacGinitie et MacGinitie, 1989)	Lecture de paragraphes, questions à choix multiple sur les passages, nécessitent une compréhension de surface du texte et des inférences (allant au-delà de ce qui est énoncé dans le texte)
Compréhension de phrases (Wagner *et al*., 2010)	Les participant·es lisent le plus de phrases possible et indiquent si celles-ci sont logiques ou non en 3 minutes.
Passage comprehension subtest du Woodcock-Johnson III (WJ-III) (Woodcock *et al*., 2001)	Lecture de courts passages, identification de mots-clés. La difficulté des éléments s'accroît progressivement (par augmentation de la longueur du passage, la rareté du vocabulaire et la complexité des indices syntaxiques et sémantiques).
Le sous-test de compréhension de passage du WRMT-R (Woodcock, 1987)	Mesure la capacité à comprendre des mots et des phrases par le biais d'images et autres indices contextuels
TABE (CTB/McGraw-Hill, 2008)	Test de compréhension conçu spécifiquement pour les adultes inscrits à des programmes d'alphabétisation :

	lecture de brefs passages et questions à choix multiples
ECOSSE (Lecocq, 1996)	Les sujets doivent lire une phrase et l'associer à une des 4 images présentées. Il s'agit de tests de compréhension avec accent sur les structures syntaxiques.
Lecture de mots et compréhension, revisée (Khomsi, 1990)	Épreuve composée de 20 planches où se trouvent 4 images et une phrase. Le sujet doit lire la phrase puis désigner l'image qui lui correspond.

Note. Seules les épreuves réputées mesurer la compréhension en lecture, et non le décodage, sont recensées.

Les épreuves, standardisées ou non, varient sur le plan du genre de texte qu'elles utilisent. La littérature sur les genres de textes décrit bien combien l'adoption d'un genre déclenche des a priori chez le lecteur ou la lectrice qui guide sa lecture (Bronckart, 1996). Les mesures sur la compréhension en lecture n'impliquent que rarement la lecture de véritables textes : elles se font généralement sur la base de *textoïdes* (Graesser *et al.*, 1997), soit des textes manipulés, dont les variables externes aux objectifs de recherche ont été contrôlées (fréquence lexicale, longueur des phrases, structures syntaxiques). Ces textoïdes possèdent en outre une faible validité écologique, et sont souvent peu intéressants ou même incohérents, ce qui peut présenter un réel risque méthodologique : celui de générer chez les sujets des représentations erronées ou des stratégies de lecture non naturelles (Graesser *et al.*, 1997, p. 166).

Les questions à choix de réponses, portant souvent sur des passages très courts, échouent à mesurer des informations de haut niveau (Keenan *et al.*, 2008). Les

épreuves de ce type sont en effet réputées mesurer des habiletés de bas niveau, soit des éléments explicitement présents dans les textes (Rupp *et al.*, 2016). Ces tests font par ailleurs partie de trousses d'évaluation du langage ou d'habiletés cognitives générales qui ne permettent pas de mesurer la compréhension à son niveau culminant, soit le modèle de situation (Kintsch, 1998) ou la structure créée (Gernsbacher, 1997).

Il appert donc que les épreuves qui mesurent les performances en compréhension de la lecture ne sont pas nécessairement arrimées à la recherche ni même aux éléments de convergence au sein des travaux sur la compréhension. Elles sont éloignées des modèles théoriques de la compréhension langagière, dans la mesure où les trois différents niveaux de la compréhension sont peu pris en compte (Farrall, 2012)[60].

[60] Des critiques d'autres ordres peuvent également être mises de l'avant par rapport à l'évaluation de la compréhension en lecture. Mentionnons crucialement ce parti-pris pour la lecture dite « littéraire », « experte », ou « savante », plutôt que pour la « lecture subjective », issue du modèle du sujet lecteur, qui a notamment profondément influencé la didactique de la littérature dans les 30 dernières années, avec le souci de redonner un sens – considéré comme perdu – à l'enseignement de la littérature (Daunay, 2007).
La mesure des performances en lecture au sein de la psychologie et des bilans orthophoniques ne tient généralement pas compte de la subjectivité propre à la lectrice ou au lecteur, et considère le produit de la lecture comme une entité théorique, rigide, et somme toute, assez élitiste, et le lecteur ou la lectrice, comme une entité abstraite, idéalisée. Cette position est clairement issue d'une approche dite « savante » de la lecture. Elle relève aussi accessoirement d'une vision de l'acte de lire qui met de l'avant les codes occidentaux et occulte les autres (Depallens, 2020). Le modèle de la lecture savante constitue une approche de la lecture dominante dans le champ de la psycholinguistique et dans les bilans orthophoniques. La lecture scolaire n'échappe pas non plus à cette tendance : elle est encore issue de pratiques de lecture savante dont la pertinence didactique est assez peu remise en question (pour une discussion complète à ce sujet, voir Depallens, 2020). Malgré ses évidentes limites, ce modèle semble d'une part mieux à même de capturer la variabilité interindividuelle, et d'autre part, permet la comparaison avec d'autres études dans des domaines connexes.

4.6 Synthèse du cadre théorique sur la compréhension en lecture

La compréhension en lecture est fonction de l'intégration explicite et implicite de processus de bas et de haut niveaux (Cain et Oakhill, 2007; Lonigan, 2015; Tighe *et al.*, 2018). L'étude de la compréhension en lecture et des mécanismes de bas et de haut niveaux qui la sous-tendent se caractérise par une prolifération de modèles et de domaines s'étant approprié son explication, notamment les sciences cognitives (Gernsbacher, 1991; McNamara et Magliano, 2009).

Parmi les modèles multicomponentiels de la compréhension en lecture, le modèle le plus basique est le modèle simple de la lecture (*Simple View of Reading*) (Gough et Tunmer, 1986), qui considère le construit comme le produit du décodage, habileté de bas niveau par excellence, et de la compréhension orale. Le modèle CCR (*Cognitive Components-Resource Model of Reading*) (Hannon, 2012, 2016; Hannon et Daneman, 2001) propose un ensemble de relations entre les processus de bas niveau (le décodage), les processus de niveau supérieur, soit l'intégration d'informations explicites et implicites du texte aux informations textuelles et aux connaissances préalables, et les ressources cognitives limitées qui sont partagées par de nombreuses fonctions cognitives (par exemple, la mémoire de travail).

De plus, nous avons mis en lumière le fait que la compréhension mobilise des processus (accès au lexique, raisonnement, inférences, activation, assemblage) et des représentations (structure de surface, base propositionnelle, modèle de situation),

au même titre que de nombreuses habiletés cognitives. Les modèles de Kintsch et van Dijk (1978) et de (1995) sont particulièrement éclairants en ce sens. Dans le cadre de la présente thèse, la performance en compréhension en lecture telle que nous la concevons correspond à une finalité, soit à la réussite de l'acte de comprendre un texte écrit. Les notions d'inférences (Gernsbacher, 1991), de cohérence (Oakhill et Garnham, 1988) et de modèle mental sont au cœur de ce que nous concevons comme faisant partie de l'acte de lire et de comprendre un texte.

Finalement, nous avons montré que les études empiriques sur la compréhension, en particulier celles qui cherchent à montrer les liens qu'elle entretient avec les habiletés métalinguistiques, que nous avons présentées dans le cadre théorique, sont finalement assez peu à même de représenter la complexité représentationnelle de la compréhension, en particulier de son niveau le plus élaboré : le modèle de situation. Nous proposerons, dans la partie II, de surmonter cet écueil méthodologique majeur par la présentation de deux épreuves de compréhension inédites.

L'idée qu'un principe [celui de la relativité galiléenne] ayant une généralité tellement large puisse valoir dans un domaine de la physique [la mécanique], mais en même temps ne pas valoir dans un autre domaine [l'électromagnétisme] me frappa comme à priori assez peu probable.

— Albert Einstein

CHAPITRE V

SYNTHÈSE

Dans la partie 1, que le présent chapitre clôt, quatre construits théoriques ont été présentés : les habiletés métalinguistiques, la conscience morphologique, la conscience syntaxique et la compréhension en lecture. Dans ce chapitre, nous ferons une synthèse du cadre théorique afin de présenter le raisonnement donnant lieu à nos questions de recherche et d'asseoir notre problématique, au fondement de notre cadre méthodologique, qui sera présenté dans la partie 2.

Les différents construits de la cognition humaine, dont ceux que nous avons exposés dans les précédents chapitres, soit la conscience morphologique (chapitre II), la conscience syntaxique (chapitre III) et la compréhension en lecture (chapitre IV), sont issus de l'analyse de comportements complexes et variés. Nos hypothèses interrogent les modèles sur les interactions entre mécanismes et composantes régissant les différents comportements afin d'en arriver à une compréhension plus fine des habiletés langagières, notamment en lecture. Plus précisément, nous cherchons à circonscrire les enjeux liés au choix des outils de mesure des habiletés métalinguistiques afin de faire la lumière sur la relation entre les HmL et la compréhension en lecture.

Nous avons donc présenté deux construits métalinguistiques importants en raison de leur faculté de prédire les performances en compréhension en lecture chez des enfants d'âge scolaire, soit la conscience morphologique et la conscience syntaxique (Carlisle, 2000; Deacon et Kirby, 2004; Demont et Gombert, 1996; Nagy *et al.*, 2006; Tong *et al.*, 2014). Ces deux construits représentent les habiletés associées à la « conscience grammaticale » (Blakmore et Pratt, 1997; Bowey, 1994).

Notre revue de la littérature nous a amenée à déterminer deux éléments constitutifs du « métalinguistique » qui se retrouvent dans la typologie des épreuves transformationnelles mesurant les HmL : d'abord, la *langue comme objet de l'attention volontaire*, ensuite, le *contrôle et la manipulation* des éléments langagiers. Il s'agit

globalement des deux dimensions représentées dans le modèle de Bialystok et Ryan (1985), fruit d'une reconceptualisation des construits effectuée à travers une analyse des demandes procédurales différenciées que mobilisent les épreuves mesurant les HmL.

Dans le prolongement de ces efforts de reconceptualisation des HmL, notre revue méthodologique de la conscience morphologique et de la conscience syntaxique a fait émerger l'hypothèse selon laquelle les épreuves élaborées pour mesurer la conscience morphologique et la conscience syntaxique font fonctionner deux types de variables :

- le type de l'unité traitée, variable linguistique, représentationnelle : morphème, groupe syntaxique ;
- l'opération demandée, variable cognitive, processuelle : ajout et suppression.

La présente thèse aborde les performances en lecture à travers le prisme des notions d'unités (*language-specific*) et d'opérations (*domain-general*), par la présentation d'une forme de parallélisme sur le plan des processus et des représentations entre les différents construits traités dans les chapitres du cadre théorique : conscience morphologique, conscience syntaxique et compréhension en lecture.

Notre premier objectif de recherche consiste à vérifier si les habiletés métalinguistiques se réfèrent à ces deux dimensions, une première de nature

linguistique, représentant le type d'unités traitées, la seconde, cognitive, représentant le type d'opérations mentales effectuées. Nous cherchons ainsi à répondre à la question de recherche suivante :

- *Les mesures traditionnelles des habiletés en conscience morphologique et en conscience syntaxique peuvent-elles se scinder en deux constituants, l'un linguistique, l'autre cognitif, prenant la forme de variables latentes ?*

Pour ce faire, nous utilisons des épreuves non verbales, qui sont, à l'instar des matrices de Raven (Fajgelj *et al.*, 2010; Raven, 1938), des mesures du raisonnement visuospatial. Ces épreuves, dont nous faisons l'hypothèse qu'elles représentent les opérations cognitives de suppression et d'ajout avec des unités non linguistiques, seront présentées dans la partie 6.4.3. Elles nous permettent, d'une part, de tester l'hypothèse de l'existence de ces opérations mentales minimales comme sources de variation au sein des épreuves, et d'autre part, de nous assurer que les performances des sujets ne sont pas imputables à des différences d'intelligence (Bialystok *et al.*, 2014). Dans la même logique, nous cherchons à construire des tâches de CM (6.4.1) et de CS (6.4.2) qui diffèrent seulement sur les plans qui nous intéressent, soit les unités manipulées – qui portent théoriquement la même valeur de difficulté dans le modèle de Bialystok et Ryan (1985) –, les opérations cognitives (suppression, ajout), et éventuellement, le contrôle attentionnel et l'analyse des connaissances (implicite, explicite).

Les variables latentes dont nous pensons qu'elles circonscrivent deux dimensions susceptibles d'induire des différences dans la part de variance explicative des performances en compréhension en lecture seront mises en relation avec les performances générales en compréhension en lecture. C'est dans l'aboutissement de ce raisonnement que prend forme notre deuxième question de recherche :

- *Dans quelle mesure les variables latentes construites – d'ordres linguistique et cognitif – contribuent-elles aux performances en compréhension en lecture?*

Cette question peut sembler ancienne, classique. Elle est pourtant novatrice, dans la présente thèse, puisqu'elle se pose après la recombinaison des variables indépendantes en dimensions cognitive et linguistique.

La thèse cherche également à repenser les exercices et les activités en classe autour du fonctionnement de la langue, au travers des dimensions dévoilées par la théorie. La question de recherche associée à ce troisième objectif est la suivante :

- *Comment peut-on opérer le passage entre tâches métalinguistiques et activités grammaticales?*

Cette question de recherche s'inscrit dans le champ de la didactique du français. Nous cherchons à mieux définir les savoirs mobilisés de manière à penser une transposition

didactique des épreuves métalinguistiques dans la classe de français du 2e cycle du primaire.

PARTIE 2 CADRE MÉTHODOLOGIQUE

What we observe is not nature itself, but nature exposed to our method of questioning.

— Werner Heisenberg, *Physics and Philosophy: The Revolution in Modern Science*

CHAPITRE VI

INSTRUMENTS DE MESURE

Dans ce chapitre, nous présenterons les épreuves utilisées pour répondre à nos questions de recherche. Trois grands construits théoriques sont mesurés : la conscience morphologique, la conscience syntaxique et la compréhension en lecture. Des épreuves matricielles non verbales, réputées impliquer des opérations de suppression et d'addition, sont aussi présentées. La partie 6.1 est dédiée à une description synthétique du fonctionnement des épreuves, et la partie 6.2 détaille le matériel et le déroulement relatifs à chacun des instruments de mesure utilisés.

6.1 Description des épreuves

Le schéma est le suivant : en écho aux épreuves transformationnelles en conscience phonologique (Gombert, 2002), citées par Demont *et al.* (2006), d'élimination, d'inversion (Demont et Gombert, 1996), de remplacement (Armand, 2005) et d'ajout (Stanovich *et al.*, 1984) de phonèmes, des tâches impliquant des unités morphologiques et syntaxiques sont construites. Celles-ci sont inspirées des épreuves de production de mots dérivés, d'extraction de la base dans des mots affixés et de production de néologismes (Reder *et al.*, 2013). Le format de ces épreuves est analogue à celui qui a cours dans la littérature sur la conscience morphologique et la conscience syntaxique ; elles reproduisent effectivement les types de transformation propres aux épreuves de conscience phonologique avec des unités grammaticales allant de l'affixe au syntagme. Les épreuves créées spécifiquement pour la présente thèse ont aussi été pensées en fonction des opérations d'addition et suppression qu'elles mettent en œuvre.

Certaines variables de contrôle, soit l'âge et des habiletés non verbales – mesures du facteur g[61] (Cattell *et al.*, 1973a) –, sont aussi prises en compte. Ces variables, impliquant un matériau non linguistique, et qui sont représentées par des épreuves

[61] Le facteur g (pour *general factor*) est un construit mis en évidence par Spearman (1904) à l'aide d'analyses factorielles. Il est utilisé régulièrement en psychométrie pour faire référence aux corrélations constantes relevées entre les mesures aux tests d'aptitudes cognitives et d'intelligence. Les matrices de Raven sont réputées depuis longtemps être des mesures « pures » du facteur g (Spearman, 1946, p. 127). La réalité psychologique de ce facteur est débattue, au regard, notamment, de son relatif isomorphisme avec d'autres construits cognitifs importants (cf. Demetriou *et al.*, 2018).

impliquant des matrices d'images, sont considérées impliquer les mêmes transformations d'addition et de suppression que celles mobilisées dans les épreuves métalinguistiques (Mackintosh et Bennett, 2005). Ainsi, au total, deux types d'épreuves, certaines mobilisant une opération de suppression, d'autres mobilisant une opération d'ajout, seront appliqués à deux unités linguistiques – morphème et groupe syntaxique – et à une unité non linguistique. Dans l'optique de construire des variables dépendantes avec un grain le plus fin possible, toutes les tâches sont réalisées à l'ordinateur, à l'exception de la tâche de compréhension en lecture avec texte long, effectuée sur papier puis recodée. Cette démarche permet d'obtenir de l'information sur le décours temporel des processus, en plus de la qualité de la réponse. La modalité de passation des épreuves est uniformément écrite. La modalité écrite a l'avantage de ne pas induire de surcharge de la mémoire de travail, contrairement à ce qui peut être le cas dans les épreuves à l'oral (Nation et Snowling, 2000). Celle-ci présente toutefois le désavantage d'être un facteur de confusion avec les habiletés en décodage.

Le tableau 6.1 donne une vue d'ensemble des instruments de mesure utilisés en fonction des construits théoriques qu'ils représentent. Les tâches de conscience morphologique sont réalisées à l'aide du logiciel Logix© (Foucambert et Bourdages, 2017). Les tâches de conscience syntaxique, les épreuves matricielles avec matériau non linguistique et tirées du CFT1-R et la tâche de compréhension de courts textes ont toutes été adaptées et codées dans *Psychopy* (Peirce *et al.*, 2019).

Tableau 6.1 Vue synthétique des construits théoriques étudiés, des épreuves servant à les mesurer et des logiciels de la collecte des données

Conscience morphologique 6 épreuves sur les « mots de même famille » dans Logix © (Foucambert et Bourdages, 2017)
Conscience syntaxique 3 épreuves de réplication de l'erreur dans Psychopy (Peirce *et al.*, 2019)
Contrôle/épreuves matricielles avec matériau non linguistique 2 épreuves matricielles dans Psychopy (Peirce *et al.*, 2019)
Compréhension en lecture 1 épreuve de compréhension en lecture avec petits textes dans Psychopy (Peirce *et al.*, 2019) et 1 épreuve de compréhension en lecture avec long texte sur papier

Les instruments servant à mesurer la conscience morphologique (6.4.1), les habiletés métasyntaxiques (6.4.2) et la compréhension en lecture (6.4.4), soit nos trois construits principaux, seront d'abord présentés (voir la liste des épreuves dans le tableau 6.2). Seront aussi décrites nos mesures de contrôle, c'est-à-dire les épreuves matricielles et le formulaire de consentement, comprenant les données sociodémographiques (6.4.3). Le déroulement général de la collecte est sensiblement le même pour l'ensemble des classes sondées (cf. le tableau 6.2). Toutes les épreuves sont originales : elles sont inspirées des épreuves traditionnellement utilisées dans les travaux psycholinguistiques, mais ont été construites afin de répondre spécifiquement à nos questions de recherche, en contrôlant un maximum de facteurs exogènes.

Tableau 6.2 Déroulement de la collecte

Première séance

1. Épreuve de paires de mots (ou de jugement de relations de mots)
2. Épreuve d'extraction de la base
3. Épreuve de dérivation

Deuxième séance

4. Épreuve de décomposition
5. Épreuve de choix d'affixes dérivationnels dans des pseudomots
6. Épreuve de choix d'affixes flexionnels dans des pseudomots

Troisième séance

7. Matrices
8. Classification
9. Épreuve de compréhension en lecture sur courts textes

Quatrième-cinquième séance

10. Remplacement
11. Addition suppression
12. Déplacement

Cinquième-sixième séance

13. Épreuve de compréhension en lecture sur texte long (format papier)

6.2 Instruments de mesure

Les épreuves permettant de mesurer les différents construits sont décrites dans cette section, chaque fois en deux parties : déroulement et matériel.

6.2.1 Épreuves de conscience morphologique

Les nombreuses différences relevées dans les épreuves de conscience morphologique ont été abordées et reconceptualisées dans la section 2.2. L'élaboration de cette typologie d'épreuves nous a permis, dans la présente étude, de ramener les facteurs de confusion à un minimum[62].

Tout d'abord, dans la mesure où les connaissances du vocabulaire ne sont pas le construit qui doit être mesuré dans notre expérimentation sur la conscience morphologique, la fréquence des items a été contrôlée en n'admettant que des lemmes *très fréquents*, *fréquents* ou *assez fréquents*, selon la nomenclature du logiciel Antidote© (version 9.2)[63]. Nous avons aussi contrôlé les items (bases et affixes) sur le plan de la transparence, puisque les mots phonologiquement transparents, c'est-à-dire ceux qui sont représentés par une même suite de phonèmes dans la prononciation du mot dérivé, comme *couture* dans *couturier*, sont reconnus et lus plus vite que les mots moins transparents ou opaques (Carlisle et Stone, 2005). Ainsi, dans les épreuves que nous avons élaborées, les transformations que l'élève doit effectuer pour arriver aux bonnes réponses n'entrainent la plupart du temps aucun changement phonologique, et des changements orthographiques minimaux (de l'ordre de deux graphèmes au maximum).

---

[62] La description de ces épreuves est notamment parue dans Bourdages et Foucambert (2018). Certaines parties du texte ont été reprises intégralement ici.

[63] Antidote© est un dictionnaire bien connu et constitue aussi une base de données en français. Le corpus rassemble plus de 4 milliards de mots.

La conception très stricte de notre matériel expérimental a inévitablement été aussi contrainte par les règles orthographiques du français. Les épreuves ont été construites avec le souci de contrebalancer les différents formats d'items – en délimitant la nature des bases et des affixes : par exemple, une base nominale où l'objectif est d'apposer un suffixe formant des adjectifs correspond à un format (N>Adj). Toutes les réponses obtenues ont été codées selon leur qualité et ramenées à un score binaire (0-1). Nous présentons ci-après les six épreuves utilisées dans le cadre de cette recherche.

Il est à noter que les épreuves varient tout de même sur certains plans, notamment en ce qui a trait à la présence de pseudomots et à la présence d'un contexte linguistique au sein des items. La présence de pseudomots et l'absence d'indices contextuels sont réputés bloquer le recours des participant·es à leurs connaissances lexicales, en les contraignant à véritablement porter leur attention sur la structure morphémique en jeu (Tighe *et al.*, 2018; Tighe et Schatschneider, 2015)[64].

Les items de chacune des six épreuves de conscience morphologique sont présentés aléatoirement aux sujets dans le logiciel Logix© (Foucambert et Bourdages, 2017). Seuls les items d'entrainement ont été présentés dans un ordre préétabli afin de faciliter la compréhension des objectifs pour les sujets et la présentation de celles-ci pour l'expérimentatrice (voir le protocole de passation à l'annexe A). Les items des épreuves de conscience morphologique figurent à l'annexe F.

[64] Tighe *et al.* (2018) ont d'ailleurs dû modifier leur modèle de AFC à six facteurs et faire l'hypothèse de la contribution d'une tâche de CM impliquant des indices contextuels à la variable latente représentant le vocabulaire afin de représenter cette source de variation.

6.2.1.1 Paires de mots

Dans cette épreuve, l'élève doit prendre une décision sur la relation sémantique entre les deux mots qui lui sont présentés (Allard, 2008; Carlisle et Nomanbhoy, 1993; Colé *et al.*, 2004; Duncan *et al.*, 2009; Goodwin *et al.*, 2017; Roy et Labelle, 2007; Spencer, Mercedes *et al.*, 2015).

Déroulement

Lors de la phase de pratique, l'expérimentatrice s'assure que les sujets ont compris le concept de « famille de mots », qui figure dans les consignes, à la suite de la présentation de trois types de paires (deux vraies, une fausse), avec rétroaction. Les sujets doivent appuyer sur O, pour oui, si les deux mots présentés sont de la même famille, comme c'est le cas pour la paire *école-écolier*. Il leur faut appuyer sur N, pour non, lorsque les deux mots à l'écran ne sont pas reliés sémantiquement, comme pour *bouche-boucher*.

Matériel

Le matériel est constitué de douze paires de mots ayant une base morphologique commune et de douze autres paires de mots reliés uniquement formellement (phonologiquement et orthographiquement), et qui sont pseudosuffixés, pour un total de 24 items. Les items de la tâche sont en partie tirés des études de Duncan *et al.* (2009) et de Allard (2008).

Parmi les douze vraies paires, six sont sous le format « Verbe, Nom (Base + suffixe) » (ex. *glisser, glissade*) et six autres sont sous le format « Nom, Nom (Base + suffixe) » (ex. *couture, couturier*). Les fausses paires sont construites sur le même schème (ex. *saler, salade ; colle, collier*). Lorsque certains items avaient des homophones, nous nous sommes assurée que l'item lexical choisi dénotait le lemme le plus fréquent dans Antidote (version 9.2) (ex. le nom « cour », est plus fréquent que le verbe « courir »). Les paires de mots impliquent des changements phonologique et orthographique minimaux à l'ajout du suffixe.

La fréquence a aussi été contrôlée. Ainsi, tous les mots sont considérés comme fréquents ou très fréquents dans Antidote (version 9.2) (voir l'indice de fréquence dans le tableau à l'annexe F). Nous avons aussi pris soin de tenir compte du sens dénoté par chacun des suffixes. Les paires de suffixes sont donc presque équivalentes (à l'exception de feuillage, 'ensemble de X' et village, 'lieu', qui ont été laissés tels quels pour éviter des problèmes de fréquence). Les cinq suffixes choisis sont présentés dans deux paires de mots. Il s'agit des suffixes –ade (V>N), –eur (V>N), –aison (V>N), –age (N>N), –ier (N>N) et –et (N>N). Il est à noter que les fausses paires de mots ont été construites en ayant en tête les suffixes correspondant aux vraies paires de mots (ex : guerre, guerrier; chant, chantier), mais qu'une équivalence parfaite n'a pu être atteinte. Les critères qui ont présidé à la construction des paires de mots sont les suivants : présence d'un pseudosuffixe, fréquence des lemmes, format (V>N ou N>N), aucun changement phonologique à l'ajout du suffixe et aucun changement orthographique à l'ajout du suffixe.

Les items ainsi que les caractéristiques ayant régi leur contre-balancement et leur inclusion dans l'épreuve sont présentés à l'annexe F.

6.2.1.2 Extraction de la base

L'épreuve d'extraction de la base évalue l'habileté à analyser les constituants morphémiques des mots en retrouvant leurs bases.

Déroulement

L'expérimentatrice s'assure à nouveau que les sujets ont compris le concept de « famille de mots ». Il leur est demandé de trouver un mot plus petit à l'intérieur de chaque mot présenté. Le mot repéré et écrit par le sujet doit être de la même famille que le mot présenté. Lorsque le mot patinage est présenté, par exemple, il faut écrire le mot *patin*, *patine*, ou *patiner*.

Matériel

Cette tâche avec mots et mots pseudoaffixés, comme *fillette* et *alouette*, figurait initialement dans Rubin (1988), puis a été reprise par Colé *et al.* (2004) et par Reder *et al.* (2013), qui y ont ajouté des pseudomots affixés, comme *tassette* et *dégarer*. Elle est constituée de 24 items, créés à partir de trois schèmes : d'abord, le schème base + affixe, afin de créer huit pseudomots, comme *tassette* ou *bruitet* ; ensuite le schème base + affixe pour former huit mots existant bel et bien, comme *fillette* ou *garçonnet* ; et enfin, le schème de mots pseudoaffixés, sur la base de *alouette* ou *chalet*. Tous les

items préfixés, au nombre de douze, sont des verbes, et tous les items suffixés, aussi au nombre de douze, sont des noms. La fréquence a été contrôlée ; ainsi, tous les items sont fréquents selon Antidote © (voir l'annexe F pour les détails).

L'épreuve nécessite un recodage, afin d'inclure les items mal orthographiés, mais dont l'identification de la base a été effectivement accomplie.

6.2.1.3 Dérivation

Cette tâche est régulièrement mobilisée dans la mesure de la conscience morphologique (voir notamment Apel *et al.* (2013); Carlisle (2000); Casalis *et al.* (2004); Tighe et Binder (2015)). Dans le cadre de cette épreuve, qui rappelle le protocole expérimental du *wug-test* (Berko, 1958), le sujet doit compléter une série de phrases en ajoutant un suffixe dérivationnel ou flexionnel pour chaque mot fourni (Carlisle, 1995; Tighe *et al.*, 2018).

Déroulement

Pour chaque item, on présente au sujet un mot, qui est suivi d'une phrase dans laquelle il manque un mot, comme dans l'exemple suivant : *garage. Gaston est un* _____. [*garagiste*]. On demande aux participant·es de compléter la phrase en écrivant un mot de la même famille que celui qui est en haut de l'écran, et qui rend la phrase correcte.

Matériel

Notre épreuve comporte un total de 24 items, qui figurent à l'annexe F. Le matériel pour la partie dérivationnelle comprend 16 phrases à trous, avec un mot à dériver pour chacune. En tout, huit suffixes dérivationnels sont élicités (deux fois pour chacun d'eux) : -ité (ADJ>N), -esse (ADJ>N), -ier (N>N), -ie (N>N), -ieux (N<ADJ), -u (N<ADJ), -able (V>ADJ), -atif (V>ADJ). Les items se déclinent en quatre formats Base + Suffixe : ADJ>N, N>N, N<ADJ et V>ADJ (quatre items pour chaque format).

Le matériel pour la partie flexionnelle comprend huit phrases à trous, avec un mot à décomposer pour chacune. Deux types de flexions sont envisagées : de genre et de personne. Pour des questions de différenciation phonologique, nous n'avons gardé que des verbes conjugués à la première personne du pluriel, à l'image de cet exemple : *écouter. Nous* _____ [*écoutons*]. Pour ce qui est des flexions en genre à décomposer, il s'agit de conjuguer au masculin ce qui était au féminin (ex : chanteuse-chanteur). Quatre types de suffixes sont présentés dans quatre mots, soit 2 adjectifs et deux noms : -esse (N), ère (N), -euse (Adj) et-ive (Adj).

La fréquence a été systématiquement contrôlée. Ainsi, tous les mots sont considérés comme fréquents ou très fréquents dans Antidote (version 9.2). Les dérivations à effectuer (l'ajout des suffixes) n'entrainent aucun changement phonologique, et quelques fois, un changement orthographique minimal.

Ici encore, les items sont recodés afin d'inclure toutes les bonnes réponses, notamment les items mal orthographiés dans lesquels la suppression du morphème est réussie.

6.2.1.4 Décomposition

L'épreuve de décomposition (Carlisle, 2000; Kieffer *et al.*, 2016; Tighe et Binder, 2015) se déroule sur le même modèle que l'épreuve de dérivation, et comporte les mêmes consignes.

Déroulement

Un mot est présenté, suivi d'une phrase à compléter, comme dans l'exemple qui suit : *courageux. Mario a beaucoup de* _____ [*courage*]. On demande aux participant·es de compléter la phrase en écrivant un mot de la même famille que celui qui est en haut de l'écran, et qui rend la phrase correcte.

Matériel

L'épreuve comporte au total 24 items (voir l'annexe F). Le matériel pour les items de morphologie dérivationnelle comprend 16 phrases à trous, avec un mot à décomposer pour chacune. En tout, huit suffixes dérivationnels doivent être supprimés afin de retrouver la base (deux fois pour chacun d'eux) : -eur (ADJ>N), -ise (ADJ>N), -age (N>N), -aire (N>N), -eux (N>ADJ), -ique (N>ADJ), - (V>ADJ), -atif (V>ADJ). Les items

se déclinent en quatre formats Base + Suffixe : ADJ>N, N>N, N>ADJ et V>ADJ. Il y a quatre items dans chaque format, pour un total de 16 items.

Le matériel pour la partie flexionnelle comprend huit phrases à trous, avec un mot à décomposer pour chacune. Deux types de flexions sont envisagées : de genre et de personne. Pour des questions de différenciation phonologique, nous n'avons gardé que des verbes conjugués à la première personne du pluriel pour les flexions de personne. Les 4 phrases de ce type commencent toujours par « Il faut » (ex. chassons. Il faut _____. [*chasser*])). Pour ce qui est des flexions en genre à décomposer, il s'agit de conjuguer au masculin ce qui apparait au féminin (ex. *fatigante. Mon chien est un peu _____.* [*fatigant*]), en concordance avec l'épreuve de dérivation. Quatre types de suffixes sont présentés dans quatre mots, soit deux adjectifs et deux noms : -esse (N), ère (N), -euse (Adj) et -ive (Adj).

La fréquence a été contrôlée. Ainsi, tous les mots sont considérés comme fréquents ou très fréquents dans Antidote (version 9.2). Les dérivations à effectuer (l'ajout des suffixes) n'entrainent ni changement phonologique ni changement orthographique.

6.2.1.5 Choix de pseudomots suffixés (morphèmes dérivationnels)

Dans cette tâche, les participant·es doivent compléter une phrase en sélectionnant le pseudomot qui possède le suffixe dérivationnel approprié (Allard, 2008; Kieffer *et al.*, 2016; Nagy *et al.*, 2006).

Déroulement

Dans chaque phrase présentée, un mot est manquant. On demande alors aux sujets de choisir la meilleure réponse parmi les trois choix de pseudomots : ex. *Hubert et son papa vont à la* _____ . [*louviterie ; louvitir ; louvitable*].

Matériel

Les suffixes qui ont été choisis pour construire la tâche sont tous exclusivement d'une catégorie grammaticale, c'est-à-dire qu'ils ne peuvent être utilisés que pour former des mots d'une seule catégorie grammaticale – adjectivale, verbale ou nominale[65], dans ce cas-ci (selon le *Dictionnaire des suffixes du français le Robert*). Les items sont tous formés à partir d'une pseudobase et du suffixe -ir, dans le cas des pseudoverbes, des suffixes -able, -ant, -al, -esque et -ible dans le cas des pseudoadjectifs, et des suffixes -age, -ade, -eau, -ure et -ité dans le cas des pseudonoms. L'épreuve comporte 15 items qui, pour rappel, peuvent être consultés dans l'annexe F.

[65] Il est à noter cependant que le suffixe -ant peut à l'occasion être utilisé pour former des noms qui ont tous la propriété de pouvoir être utilisés comme adjectifs (ex. *fatigant*). Le choix ou non du suffixe -ant à l'intérieur de pseudomots ne pose donc pas problème dans notre corpus, même s'il n'est pas aussi exclusivement adjectival que les autres suffixes. C'est sur cette base qu'il a été inclus dans l'élaboration de nos items.

Les suffixes choisis sont tous exclusivement d'une catégorie, c'est-à-dire qu'ils ne peuvent être utilisés que pour former des mots d'une même nature (selon le *Dictionnaire des suffixes du français le Robert*).

6.2.1.6 Choix de pseudomots suffixés (morphèmes flexionnels)

Dans cette tâche, les sujets doivent compléter chacune des phrases présentées en sélectionnant le pseudomot qui possède le suffixe flexionnel approprié (Goodwin *et al.*, 2017; Tighe *et al.*, 2018).

Déroulement

Les consignes de tâche de choix de pseudomots suffixés avec morphèmes flexionnels sont les mêmes que pour la tâche précédente, à la différence que quatre choix sont offerts, plutôt que trois.

Matériel

Dans le but de contrôler la complexité syntaxique de chacune des phrases, nous avons systématiquement construit des items où le verbe manquant est accompagné d'un autre verbe dans la phrase, comme on peut le constater dans l'exemple suivant : *Lucie m'a dit de _____ la porte. glodons ; gloder ; glodant ; glode*. Chaque phrase comporte donc deux propositions (au sens de (Riegel *et al.*, 2004))[66].

[66] Pour Riegel *et al.* (2004), et selon usage grammatical traditionnel, une proposition est faite des phrases principales et constituantes des phrases complexes. Ces phrases constituantes peuvent être « soit enchâssées dans une autre phrase, soit combinées par juxtaposition ou

6.2.1.7 Portrait global des épreuves de conscience morphologique et dimensions impliquées dans leur traitement

Comme mentionné *supra*, nous avons contrôlé certains facteurs pouvant influencer les résultats, notamment la fréquence, la transparence phonologique et les modalités de passation. Les épreuves varient toutefois sur le plan des unités qui sont mises en jeu et sur les types de manipulations mobilisées (Bourdages et Foucambert, 2018). Le tableau 6.3 donne un aperçu des tâches et de leur contrebalancement au regard des trois différentes dimensions dont nous allons évaluer le pouvoir descriptif et prédictif – l'unité, le type de contrôle et l'opération.

Tableau 6.3 Aperçu des tâches de conscience morphologique utilisées et de leur classement en fonction de deux dimensions

Épreuves	**Unité**	**Opération**
1. Paires de mots	Mots dérivés et pseudodérivés avec suffixes dérivationnels	Suppression
2. Extraction de la base	Suffixes dérivationnels	Suppression

coordination » (p. 472). Les auteurs font la remarque suivante : « Le terme de *proposition* remonte aux grammaires logiques où il désignait toute construction minimale porteuse d'un jugement : l'association d'un sujet (ce dont on dit quelque chose) et d'un prédicat (ce que l'on dit du sujet). Ainsi les grammairiens de Port-Royal (Arnauld et Lancelot, 1660/1810) analysaient la phrase *Dieu invisible a créé le monde visible* en trois propositions : Dieu a créé le monde – Dieu est invisible – Le monde est visible. Progressivement, la notion de proposition s'est confondue avec celle de phrase pour désigner l'unité syntaxique et prédicative combinant un sujet grammatical et un groupe verbal. »
Ainsi, le terme de *proposition,* au sein de la présente thèse, admet deux acceptions. Celle qui est adoptée dans le présent chapitre sur la méthode en est une grammaticale : elle recouvre les phrases principales et constituantes, qui sont le plus souvent représentées par l'introduction de verbes. Celle qui est adoptée dans le chapitre IV sur la compréhension en lecture est celle de construction minimale porteuse d'une idée, qui est issue de la logique, dont Riegel *et al.* (2004) offrent un exemple éclairant.

3. Dérivation	Morphèmes dérivationnels et flexionnels	Ajout
4. Décomposition	Morphèmes dérivationnels et flexionnels	Suppression
5. Choix de pseudomots suffixés D	Morphèmes dérivationnels	Ajout
6. Choix de pseudomots suffixés F	Morphèmes flexionnels	Ajout

6.2.2 Épreuves de conscience syntaxique

Trois épreuves permettent de mesurer les habiletés métasyntaxiques des sujets (voir le protocole de passation à l'annexe D et les items à l'annexe E). Elles ont été divisées en fonction de l'opération à mobiliser pour les réussir, soit l'*ajout-suppression*, le *remplacement* et le *déplacement*. Toutes les épreuves demandent au sujet de passer par quatre étapes :

- Identification de l'erreur dans une phrase agrammaticale ;
- Correction de l'erreur ;
- Réplication de l'erreur sur une première phrase grammaticale ;
- Réplication de l'erreur sur une deuxième phrase grammaticale .

Ainsi, les tâches de conscience syntaxique utilisées dans notre étude vont de la plus facile (localisation) à la plus difficile (réplication), selon la typologie de Gaux et Gombert (1999b).

6.2.2.1 Déroulement

Afin de déjouer la limite du temps pouvant être alloué à la collecte de données dans les classes, il a été décidé de faire se succéder les consignes, ou étapes, à appliquer sur les mêmes items. Les étapes d'identification, de correction et de réplication représentent des épreuves différentes dans certaines études (Cohen, 2011), notamment dans Gaux et Gombert (1999b), alors qu'elles sont imbriquées (*nested*) dans le cadre de notre projet[67]. Les réponses à chaque phase d'un même bloc sont ainsi fonction des réponses ayant précédé, dans la mesure où la non-identification de l'erreur mènera nécessairement à un échec sur le plan de sa correction, par exemple.

Ajout-suppression

L'épreuve fonctionne par bloc de trois phrases, en quatre temps. Premièrement, une phrase agrammaticale, dans laquelle il manque un mot ou un groupe de mots, est présentée comme dans l'exemple « le copain • ma mère est gentil ». Les sujets doivent cliquer sur le point dans la phrase où il manque un élément.

Deuxièmement, il faut corriger la phrase en écrivant un nouveau mot ou un nouveau groupe de mots. Dans notre exemple, le sujet doit écrire le mot « de ».

[67] Sur le plan statistique, cet état de choses constitue un problème de non-indépendance des informations, qui doit être envisagé au travers d'une analyse multiniveaux. En effet, le « bloc » (chaque bloc de 3 phrases, réalisées en 4 réponses successives) est un facteur aléatoire non contrôlé du côté des items. Nous y reviendrons dans le chapitre VI portant sur l'analyse des données et les résultats.

Troisièmement, une nouvelle phrase, grammaticale, cette fois, est présentée à l'écran, par exemple « le chien de Louis est doux ». Il faut cliquer sur le mot ou le groupe de mots à enlever pour reproduire l'erreur de la première phrase dans cette deuxième phrase.

Quatrièmement, cette même erreur doit être reproduite dans la troisième phrase, par exemple « ces chats roux du Portugal semblent très fatigués ». Les sujets doivent sélectionner la préposition « du ».

Remplacement

L'épreuve fonctionne par bloc de trois phrases, en quatre temps. Premièrement, le sujet doit cliquer sur le mot qui pose problème dans une phrase qui comporte une erreur, comme dans « le chat dont ta sœur aime est doux ».

Deuxièmement, la phrase doit être corrigée en écrivant un nouveau mot, soit le subordonnant « que », dans l'exemple donné.

Troisièmement, une nouvelle phrase est présentée à l'écran, par exemple « la tortue que vous observez est vieille ». Elle est bien écrite, sans erreur. Le sujet est amené à reproduire l'erreur de la première phrase dans cette deuxième phrase. Dans ce cas-ci, il doit ainsi remplacer le subordonnant « que » par le subordonnant « dont ».

Quatrièmement, il faut reproduire cette même erreur dans la troisième phrase, par exemple dans « le coiffeur bavard que ma mère adore utilise des ciseaux coupants ».

Ici encore, le sujet doit remplacer le subordonnant « que » par le subordonnant « dont ».

Déplacement

L'épreuve fonctionne par bloc de trois phrases, en quatre temps. Premièrement, la phrase agrammaticale présentée comporte un mot ou un groupe de mots qui est mal placé, comme dans l'exemple suivant : « des fleurs la femme apporte ». Il faut cliquer sur le mot ou le groupe de mots qui pose problème. Ici, il s'agit du complément direct « des fleurs ».

Deuxièmement, il faut rendre la phrase correcte en cliquant sur l'endroit dans la phrase où déplacer ce mot ou ce groupe de mots. Dans notre exemple, le sujet doit cliquer à la toute fin de la phrase, après le verbe, pour y réintroduire le complément direct.

Troisièmement, une nouvelle phrase est présentée à l'écran : « le fermier loue des tracteurs ». Elle est bien écrite, sans erreur. On demande aux sujets de cliquer sur le mot ou le groupe de mots à déplacer pour reproduire l'erreur de la première phrase dans cette deuxième phrase, puis sur l'endroit où le déplacer. Dans notre exemple, le sujet doit déplacer le complément direct « des tracteurs » au tout début de la phrase, afin de reproduire la phrase agrammaticale sur le modèle de la phrase du début : « des tracteurs le fermier loue».

Quatrièmement, les sujets doivent reproduire cette même erreur dans une troisième phrase, par exemple « un petit robot répare la voiture brisée ». Dans notre exemple, on demande ainsi au sujet de sélectionner « la voiture brisée », puis de cliquer au tout début de la phrase, afin de reproduire la phrase agrammaticale sur le modèle de la phrase du début : « la voiture brisée un petit robot répare ».

Les parties 6.2.2.2, 6.2.2.3, 6.2.2.4, 6.2.2.5 et 6.2.2.6 fournissent le détail des éléments ayant présidé à la confection du matériel expérimental des trois épreuves, donc aux phrases construites pour l'épreuve. En 6.2.2.7 se trouve une synthèse globale des épreuves de conscience syntaxique.

6.2.2.2 Éléments syntaxiques impliqués

Comme nous nous intéressons à l'étude des habiletés métasyntaxiques en général, donc à la diversité de leurs manifestations (à l'instar de Cohen (2011); Fortier (2013), notamment), des précautions ont été prises à l'étape de la constitution de nos items. Nous avons visé des éléments syntaxiques assez variés, afin de représenter le plus fidèlement possible la diversité des éléments syntaxiques de la langue, ou, à tout le moins, d'éviter la surreprésentation de certains phénomènes. En outre, nous nous sommes assurée que ces éléments n'étaient pas complètement opaques pour les élèves participant à l'étude ; nous avons vérifié que les élèves avaient déjà été en

contact avec les éléments linguistiques à manipuler dans les épreuves, par leur présence dans le plan d'études romand (PER) (CIIP, 2006)[68].

Les items diffèrent sur le plan de la fonction des éléments erronés présentés (fonction de l'élément/phénomène syntaxique erroné). Il y a chaque fois trois items par fonction représentée, pour un total de 21 items.

[68] Le plan d'études romand, ou PER, est le programme d'enseignement qui a cours en Suisse romande. À l'instar du Programme de formation de l'école québécoise, le PER spécifie un projet global de formation et d'objectifs des élèves de toute la Romandie, et ce, durant l'ensemble de leur scolarité obligatoire.

Tableau 6.4 Aperçu des caractéristiques des items dans les tâches de conscience syntaxique

Éléments internes au syntagme (mots)
- Dét (deux occurrences dans GN sujet, une occurrence dans GN CD)
- V dans prédicat à deux arguments (incluant le verbe)
- Prép dans CI du verbe
- Subordonnant dans Psub CdN

Syntagmes
- CD (3 GN)
- Sujet (3 GN)
- CdN (un GAdj, une PSub et un Gprép)

Types d'erreurs
- Élément mal placé
- Élément absent
- Élément erroné

Il est à noter que, dans le cas où l'erreur à répliquer est une suppression, plusieurs opérations ont cours pendant les trois phases. La correction de l'erreur implique forcément un ajout, alors que la réplication requiert une suppression.

6.2.2.3 Critères de construction des phrases

Certains critères ont guidé la conception de nos items, donc des phrases présentées aux sujets. Toutes les phrases sont déclaratives, à l'exception des phrases avec préposition, dans lesquelles il y a des phrases passives. Il y a aussi des phrases avec relatives, pour que le subordonnant ou la subordonnée soit identifié. Il n'y a ni phrase interrogative, ni phrase exclamative.

La fréquence lexicale a été contrôlée avec *Lextutor* (Cobb, 2013). Seuls 15 mots, sur 484, soit moins de 5 %, sont au-dessus de la barre des 10 000 mots plus fréquents. La position du mot ou groupe de mots à identifier dans chacune des phrases varie afin d'éviter la réussite des épreuves sur la base d'une analyse superficielle de la phrase, par comptage. Ainsi, chaque item est constitué d'une phrase cible B, systématiquement construite sur le modèle de la phrase cible A, mais qui inclut des groupes plus complexes, sur recommandation de Gaux et Gombert (1999b).

La complexité syntaxique a été aussi contrôlée. Il n'y a pas plus de 2 verbes, ou deux propositions, au sens de Riegel *et al.* (2004), par phrase, et un seul verbe – une seule proposition – quand il n'y a pas de phrase subordonnée. Il n'y a pas plus de trois actants[69] par phrase, et toujours dans un ordre canonique, sur recommandation de Dick *et al.* (2002).

Les éléments à manipuler au sein des phrases dans les épreuves métasyntaxiques sont les erreurs en tant que telles. Les éléments que nous avons contrôlés sont les suivants :

- Les inversions portent sur les *fonctions* (compléments, sujet) ou sur les *éléments internes* (catégories : nom, déterminant, verbe, subordonnant, etc.)

[69] Le terme « actant » désigne en linguistique le plus souvent le sujet (argument externe du verbe) et les compléments (arguments internes du verbe) (Riegel *et al.*, 2004).

- Les remplacements ne portent que sur les éléments internes (catégories : nom, déterminant, verbe, subordonnant, etc.).
- Lorsqu'il faut ajouter un élément (élément absent), on peut avoir affaire à des éléments internes (catégories : nom, déterminant, verbe, subordonnant, etc.). Il est difficile de faire une opération de remplacement sur des syntagmes entiers, car une telle action implique une forme de création de la part des sujets, le recours à la production de texte.
- Le type de contrôle (au sens de Bialystok, 2001a) diffère. La localisation de l'erreur est réputée impliquer une analyse syntaxique implicite. La correction constitue une opération légèrement plus explicite. La reproduction de l'erreur est plus complexe et explicite (sans qu'une identification des éléments syntaxiques à reproduire ne soit nécessaire).
- Les opérations mobilisées, en accord avec nos hypothèses, sont les suivantes :
    - Élément absent (opération de suppression) : 6 items
    - Élément mal placé (opération de déplacement, donc suppression et ajout) : 8 items
    - Élément erroné (opération de remplacement, donc suppression et ajout) : 4 items, opération effectuée seulement sur des mots, ou éléments intrasyntagmatiques

- Dans un même bloc, les phrases cibles A et B ont des structures identiques à la phrase modèle. Lorsque l'élément à opérationnaliser est interne et représente un mot, le mot visé est différent dans les trois phrases.

Seuls les trois premiers cas de figure sont inclus dans le matériel expérimental. Il aurait été possible de couvrir l'ensemble des quatre possibilités de types d'erreurs et de planifier l'opération ajout, avec des phrases du type 'Le maitre est contente à de son élève'. L'absence de ce type d'erreurs dans les études recensées (sauf dans Marcotte (2021)) et son lieu d'application étant limité aux mots seuls, ou éléments intrasyntagmatiques, nous a convaincue de ne pas l'inclure dans la constitution de notre corpus d'items.

6.2.2.4 Contrebalancement

Le contrebalancement des types d'opérations mobilisées n'a pu être effectué parfaitement, à cause de la nature des éléments linguistiques mobilisés (voir les fonctions représentées). Le fait que les élèves reproduisent l'erreur sur le modèle de la phrase en se basant sur la position du mot ou des mots à déplacer (donc sans analyse explicite de la structure syntaxique) a été contrôlé (Gaux et Gombert, 1999b). À l'instar de Gaux et Gombert (1999b, p.55), nous avons modifié « la nature des mots ou la composition des syntagmes […] afin d'obliger le sujet à utiliser la fonction des mots ou des syntagmes ». On propose donc deux types de phrases à reproduire : l'une (A) étant structurellement identique, et l'autre (B) comprenant des sous-groupes

en plus ou en moins (notamment par pronominalisation des GN en position sujet : « ces joueurs mécontents », par l'ajout de GAdj au sein des GN en position sujet ou des compléments, comme « ce courageux chevalier poursuit le gros dragon » et par l'ajout de Gprép : « la victoire de leur équipe »), afin de pallier la stratégie d'analyse mot à mot, ou à toute autre forme d'analyse ne relevant pas de métasyntaxe. Voici en (1) un exemple provenant de l'épreuve de déplacement :

(1) Phrase modèle : *Son frère suit <u>de natation</u> des cours.*

Phrase cible A : *Son ami veut une tasse <u>de café</u>.*

Phrase cible B *: Une étonnante élève porte une veste <u>de laine jaune</u>.*

Voici en (2) un exemple provenant de l'épreuve d'addition-suppression (intrasyntagme) :

(2) Phrase modèle : *Le lion __ le zèbre.*

Phrase cible A : *Le garçon <u>cueille</u> la rose.*

Phrase cible B : *Elle <u>adore</u> son nouveau jouet.*

6.2.2.5 Ordre de difficulté dans les tâches

Les tâches métalinguistiques requérant un niveau de contrôle attentionnel plus élevé sont réputées demander une charge cognitive plus lourde, dans la mesure où l'élève doit inhiber un ou des éléments discordants afin de produire la réponse attendue.

Ainsi, les tâches de conscience syntaxique utilisées dans notre étude vont de la plus facile (localisation) à la plus difficile (réplication), selon la typologie de (Cohen, 2011).

6.2.2.6 Limites des tâches de conscience syntaxique

La constitution des épreuves, avec le regard analytique particulièrement restrictif, sur le plan méthodologique, que nous avons souhaité adopter, a mis au jour les difficultés inhérentes à ce type d'entreprise lorsqu'il est question du langage. En effet, la conception des épreuves a été confrontée à des difficultés proprement linguistiques liées à l'idée saussurienne selon laquelle « la langue est un système où tout se tient». Ainsi, mettre la focale sur un aspect particulier du langage sans tenir compte de l'ensemble de ses aspects constitutifs est difficile, voire impossible. Les problèmes qui ont été déclenchés par cette entreprise, dans le cadre de la conception de nos épreuves, sont répertoriés ici.

Le déplacement du verbe équivaut au déplacement du CD (donc la catégorie déplacée (intra ou inter) perd de sa pertinence. Il aurait été possible d'ajouter des items dans lesquels l'erreur correspondrait à un élément absent (obligatoirement un élément interne au syntagme). L'opération demandée au sujet dans ce cas de figure aurait été un ajout. Pour assurer un certain contrebalancement, il aurait fallu inclure des items dans lesquels l'erreur correspondait à un élément en trop, pour laquelle le sujet serait amené à faire une opération de suppression. Nous avons choisi de ne pas inclure ce type d'items demandant beaucoup de créativité, d'abord parce qu'ils sont

presque introuvables dans la littérature sur la conscience syntaxique, mais aussi parce qu'ils nous semblaient difficiles à mettre en œuvre pour de jeunes élèves.

La possibilité d'analyse de la phrase à l'aide de la position du mot est réduite par l'introduction d'une deuxième phrase cible (cible B). Celle-ci nous assure aussi de la variabilité dans les résultats.

Plusieurs types de phrases ou éléments internes ne sont pas inclus dans les items, notamment, les phrases causatives – ex. « Il nous fera perdre beaucoup d'argent »–, les phrases passives – ex. « La voisine est dérangée par le bruit » – et les phrases comparatives – ex. « Mon sofa est plus doux que le tien ».

Il est à noter que contrairement à ce qui a été le cas pour l'ensemble des autres épreuves, tant celles construites dans Logix© que celles dans Psychopy©, les différents blocs des trois épreuves de conscience syntaxique n'ont pas été randomisés, étant donné la difficulté inhérente à la mise en place de ces épreuves écrites sur le plan de la programmation.

Nos trois tâches métasyntaxiques – addition-suppression, déplacement et remplacement – ne mettent pas toutes en œuvre les mêmes éléments syntaxiques, une situation inhérente à la nature des opérations qu'elles mobilisent. Notre design expérimental est donc fragile sur ce plan, puisque les différences dans les résultats

peuvent être attribuées à la nature des éléments syntaxiques rencontrés. Ce facteur de variation ne pourra être contrôlé que par un traitement statistique particulier (niché - *nested*), qui fera l'objet d'une discussion plus approfondie et d'un traitement particulier dans le chapitre VII.

6.2.2.7 Portrait global des épreuves de conscience syntaxique et dimensions impliquées dans leur traitement

Comme mentionné *supra*, nous avons contrôlé certains facteurs pouvant influencer les résultats, notamment la fréquence lexicale et les modalités de passation. Les épreuves varient toutefois sur le plan des unités mises en jeu et sur les types de manipulations mobilisées (Simard, D. *et al.*, 2017). Le tableau 6.5 donne un aperçu des tâches et de leur contrebalancement au regard des trois différentes dimensions dont nous allons évaluer le pouvoir descriptif – l'unité, le type d'analyse, le niveau de contrôle (Bialystok, 2001a; Simard, D. *et al.*, 2017) et l'opération.

Tableau 6.5 Aperçu des tâches de conscience syntaxique utilisées et de leur classement en fonction de deux dimensions.

Épreuves	**Unités**	**Opération**
Identification de l'erreur dans phrase agrammaticale M	Mot	Suppression
Correction de l'erreur dans phrase agrammaticale M	Mot	Ajout
Identification phrase grammaticale, réplication 1 M	Mot	Suppression
Correction phrase grammaticale, réplication 1 M	Mot	Ajout
Identification phrase grammaticale, réplication 2 M	Mot	Suppression
Correction phrase grammaticale, réplication 2 M	Mot	Ajout

Identification de l'erreur dans phrase agrammaticale S	Syntagme	Suppression
Correction de l'erreur dans phrase agrammaticale S	Syntagme	Ajout
Identification phrase grammaticale, réplication 1 S	Syntagme	Suppression
Correction phrase grammaticale, réplication 1 S	Syntagme	Ajout
Identification phrase grammaticale, réplication 2 S	Syntagme	Suppression
Correction phrase grammaticale, réplication 2 S	Syntagme	Ajout

Note. L'épreuve de Remplacement implique des prépositions et des subordonnants, l'épreuve d'Addition Suppression comporte des GN, des V, des prépositions et des subordonnants, et l'épreuve de Déplacement comporte des GN, des Gprép, des V, des GAdj et des P subordonnées.

6.2.3 Épreuves matricielles avec matériau non linguistique (*domain-general*)

Le *Culture Fair Intelligence Test* (Cattell *et al.*, 1973a, 1973b) vise à mesurer le fonctionnement intellectuel en réduisant au minimum l'influence de facteurs tels que le niveau de scolarisation, la langue et l'environnement culturel (dans la mouvance des tests *culture-free*) (Cattell *et al.*, 1973a, 1973b). Il s'agit d'une batterie d'épreuves non verbales dont la réussite repose sur la capacité du sujet à identifier des relations entre des figures, en quatre sous-tests. Les épreuves sont réputées être plus représentatives, ou constituer plus fortement, le facteur g d'intelligence générale fluide, plutôt que cristallisée (Cattell *et al.*, 1973a, 1973b), puisque l'intelligence établit des relations entre des éléments perçus immédiatement et sans à prioris culturels (en théorie). L'intelligence générale, ou facteur g, peut être définie comme une capacité inférentielle appliquée à des situations nouvelles (« *inferential power applied to novelty* ») (Demetriou *et al.*, 2018, p. 5).

Nous avons choisi de sélectionner deux sous-tests parmi ceux disponibles dans la batterie du CFT1-R (Weiß et Osterland, 2013), qui figurent à l'annexe H. Les sous-tests 5 et 6 du CFT1-R correspondent à des versions améliorées et plus récentes du sous-test 2 (classifications, 4 minutes) et 3 (matrices, 2 minutes ½ ou 3 minutes, en fonction des versions) de Cattell (1959/1970) pour l'édition française et révisée (1973a, b), et le sous-test 3 de Cattell *et al.* (1973a, 1973b); (Heine *et al.*, 2018)[70]. Nous avons limité à 30 secondes la présentation de chaque item, sur recommandation d'une tâche analogue dans le WISC-V (Wechsler, 2016) – voir l'annexe B pour le protocole de passation des épreuves matricielles.

6.2.3.1 Classifications

Le sous-test 5 du CFT1-R, Classifications, est une tâche de repérage d'analogies, de type « trouvez l'intrus ». Nous faisons l'hypothèse qu'ici, le sujet doit *soustraire* le trait différent de figures qui partagent par ailleurs des ressemblances, afin de réussir l'exercice.

---

[70] Cattell (1959/1970) a divisé l'intelligence générale en 2 sous-types :

- L'intelligence cristallisée : les connaissances acquises avec l'expérience, influencée par la culture et l'école.
- L'intelligence fluide : l'habileté à acquérir de nouvelles connaissances et à résoudre des problèmes, indépendante de l'apprentissage.

Il a développé un test « culture-fair » pour mesurer l'intelligence fluide, qui réduit l'influence du niveau de langue, de la culture et du niveau d'études sur les résultats. Les tests, administrés de façon individuelle ou en groupe, sont non linguistiques, c'est-à-dire que leur réussite ne requiert pas le recours au langage. Ces tests mobilisent chez les sujets leur habileté à percevoir des liens entre des formes et des figures et des stratégies de résolution de problèmes.

Déroulement

Dans cette épreuve, le sujet doit cliquer sur la seule figure qui présente une différence, qui sort du lot, parmi une série de figures.

Matériel

Cette épreuve est constituée de 15 items. Chaque item est présenté à l'écran 30 secondes au maximum. Dans chacun des items, cinq figures sont présentées au sujet. Quatre des figures présentent des similarités, mais une figure est différente des quatre autres, comme en témoigne l'exemple en 6.1. Dans cet exemple, c'est la première figure qu'il faut sélectionner.

Figure 6.1 Exemple d'item de la tâche de classifications du CFT1-R (Weiß et Osterland, 2013).

6.2.3.2 Matrices

Dans le sous-test 6 du CFT1-R, Matrices, le sujet doit compléter une matrice dans laquelle un dessin est manquant en sélectionnant une des figures présentées. Nous

faisons l'hypothèse que le sujet doit *ajouter* mentalement la figure pour réussir l'épreuve.

Cette épreuve est clairement inspirée du test d'intelligence générale classique *Raven's Colored Progressive Matrices* (Fajgelj *et al.*, 2010; Raven, 1938) (cf. figure 6.2). Les matrices progressives de Raven sont en partie de nature métacognitive (Fajgelj *et al.*, 2010). Elles sont réputées mesurer le facteur général d'intelligence, ou facteur g – tel qu'identifié à l'origine par Spearman à l'aide d'analyses factorielles avec rotations. Ainsi était établi le fait que les individus solutionnant facilement des problèmes arithmétiques ou identifiant rapidement des synonymes étaient aussi plus à même de résoudre des problèmes spatiaux et de repérer des analogies (Cattell, 1959/1970).

Les recommandations de Cattell (1959/1970) sont de ne pas présumer de l'intelligence ou de prétendre calculer le facteur g en ne faisant passer qu'une partie des sous-tests aux sujets. Les quatre soustests sont d'ailleurs relativement peu corrélés entre eux, et donc représentent des habiletés relativement différentes (Cattell, 1959/1970). Le calcul du facteur g ne constitue pas le but de la sélection de ces tests pour l'expérience. Pour rappel, ceux-ci ont été sélectionnés parce qu'ils nous semblaient pouvoir représenter et mesurer une habileté domain-general (ou dans ce cas-ci, non linguistique) qui constitue une part de la variance trouvée dans les

mesures des habiletés métasyntaxiques et de conscience morphologique, soit l'addition et la suppression.

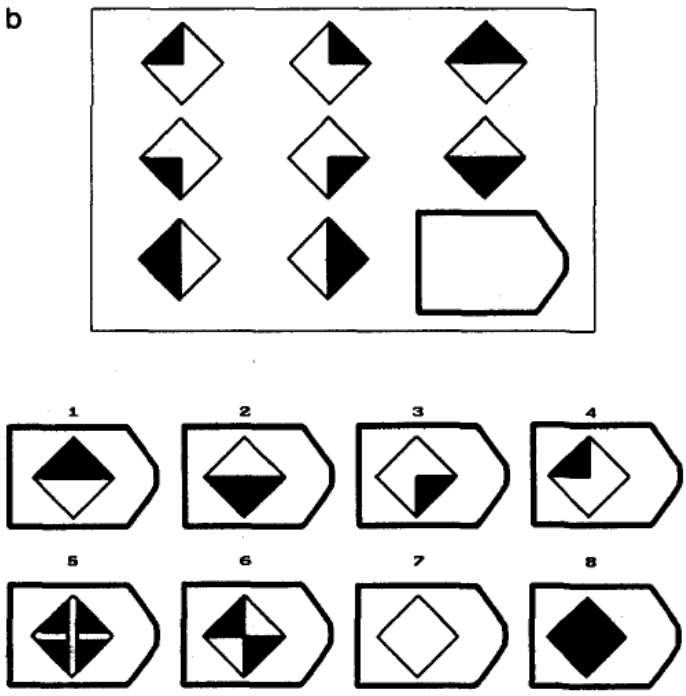


Figure 6.2 Exemple d'item du *Raven's Colored Progressive Matrices*, tel que reproduit dans Carpenter *et al.* (1990)

Déroulement

Le sujet doit compléter une matrice incomplète en cliquant sur la figure manquante. Elle ou il doit sélectionner l'une des cinq figures de la rangée de droite correspondant correctement à la case vide.

Matériel

L'épreuve est constituée de 15 items. Chaque item est présenté à l'écran 30 secondes au maximum. Dans la matrice à gauche de l'écran, il y a quatre petites cases, dont trois contiennent des figures. La figure en bas à droite est manquante. Dans le cas de l'exemple en 6.2, c'est la deuxième figure qui complète la matrice.

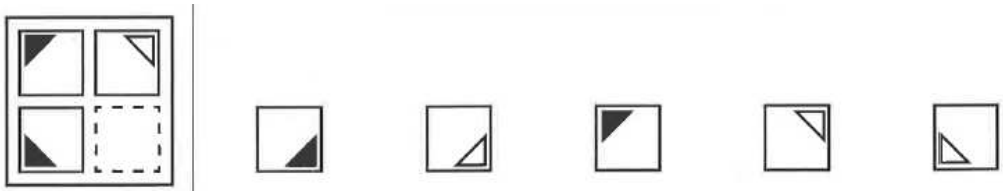


Figure 6.3 Exemple d'item de la tâche de classifications du CFT1-R (Weiß et Osterland, 2013).

Le tableau 6.6 permet de visualiser l'ensemble des épreuves transformationnelles en fonction des opérations et des unités qu'elles mobilisent.

Tableau 6.6 Aperçu des tâches métalinguistiques et transformationnelles prévues et de leur classement en fonction de deux dimensions hypothétisées

Épreuves	**Unités**	**Opérations**
Conscience morphologique		
1. Paires de mots	Mots dérivés et pseudodérivés avec suffixes dérivationnels	Suppression
2. Extraction de la base	Suffixes dérivationnels	Suppression
3. Dérivation	Morphèmes dérivationnels et flexionnels	Ajout
4. Décomposition	Morphèmes dérivationnels et flexionnels	Suppression
5. Choix de pseudomots suffixés D	Morphèmes dérivationnels	Ajout
6. Choix de pseudomots suffixés F	Morphèmes flexionnels	Ajout

Conscience syntaxique		
1. Réplication de l'erreur	Plusieurs catégories syntaxiques	Ajout et suppression
2. Correction de l'erreur	Plusieurs catégories syntaxiques	Ajout et suppression
3. Localisation de l'erreur	Plusieurs catégories syntaxiques	Ajout et suppression
Épreuves transformationnelles avec matériau non linguistique		
1. Sous-test 6 CFT1-R : matrices		Ajout
2. Sous-test 5 CFT1-R : classification (intrus)		Suppression

6.2.4 Épreuves de compréhension en lecture

Nous avons, dans le chapitre IV, démontré que la mesure de la compréhension en lecture fait débat, et que les tentatives d'évaluation ayant vu le jour dans les 50 dernières années, que ce soit dans une perspective longitudinale d'évaluation des progrès ou comparatiste (McNamara et Magliano, 2009), ne sont pas complètement arrimées aux derniers développements mis en évidence par la recherche. Les différentes approches théoriques sur la compréhension en lecture incluent toutes des niveaux de représentation – allant de la structure de surface du texte, puis la base propositionnelle du texte au modèle de situation, ou modèle mental (Gernsbacher, 1997; Kintsch, 1998; Kintsch et van Dijk, 1978). Pourtant, les épreuves permettant effectivement de juger des performances en compréhension en lecture échouent à

mesurer l'habileté à calculer les inférences et à construire un véritable modèle mental (Farrall, 2012; Hannon, 2012).

Nous proposons de surmonter cet écueil méthodologique majeur par la présentation de deux épreuves de compréhension (Foucambert, 2003) : l'une avec des questions à choix, sur textes courts ; l'autre avec des questions plus ouvertes, sur un texte long. Ces deux épreuves tentent de combler l'écart entre la recherche et l'évaluation en classe.

Nous avons conçu deux différentes épreuves de compréhension dont le protocole figure à l'annexe C et les items, à l'annexe G, qui couvrent des situations de lecture différentes et représentatives de ce qu'on attend d'un lecteur expert :

- lecture de textes courts
- lecture experte d'un texte long

Notre choix d'élaborer des épreuves basées sur des textes narratifs, ou textes qui racontent, a été motivé par l'importance de l'entrée par les genres dans l'enseignement de la langue, dans la mesure où ceux-ci président les activités liées aux pratiques langagières des élèves (Bronckart, 1996; Dolz et Gagnon, 2008). Le genre de texte constitue par ailleurs l'un des cinq niveaux de représentation construits par l'individu en cours de lecture (Graesser *et al.*, 1997).

6.2.4.1 Épreuve 10 textes courts

Notre épreuve sur textes courts est informatique, d'abord pour pouvoir recueillir les vitesses de lecture et de réponses, et également dans un souci écologique. Le matériel expérimental élaboré pour cette épreuve de compréhension de textes courts est constitué de dix textes narratifs, compris comme des textes qui racontent (conte merveilleux, récit d'aventures, fable) ou qui relatent (témoignage d'une expérience vécue), et sont inspirés d'une tâche analogue de Foucambert (2003) (cf. annexe G). Les textes, publiés en ligne et libres de droits, ont été modifiés afin de réduire l'impact de la fréquence et de la complexité syntaxique sur la compréhension (Roberts et Gibson, 2002).

Huit textes sont au présent (présent de narration). Deux textes sont au passé simple. Les dix textes présentent des niveaux comparables de difficulté : la fréquence lexicale et la complexité des phrases ont été contrôlées (longueur des phrases). L'impact de ces deux variables sur la compréhension est donc réduit. Ces deux variables ont été calculées à l'aide de Lextutor, un logiciel d'analyse automatique de textes dont la base de référence est la base de données *Lexique* (lexique.org). Par la suite, nous avons effectué une série de tests t pour échantillons uniques, avec comme valeurs de références les moyennes de chaque colonne présentées dans le tableau 6.7. Tous les tests se sont avérés non significatifs (n.s.), ce qui nous permet de conclure qu'il n'y a pas de différence significative entre les textes sur le plan de la fréquence lexicale,

de la longueur des phrases, de la longueur des textes et du nombre de mots par phrase.

Tableau 6.7 Analyse lexicométrique : portrait de la fréquence du vocabulaire et de la longueur des phrases dans l'épreuve de compréhension de textes courts selon la base Lextutor (Cobb, 2019)

Textes	% 1-1000 mots plus fréquents	%1001-2000 mots plus fréquents	% 2001-4000 mots plus fréquents	% 4001 mots plus fréquents +	% Mots hors-liste	N mots	N phrases	N moyen de mots par phrase	Écart-type
1- poisson	70,9	11,2	7,2	10,0	0,6	179	19	9,00	5,64
2- crocodile	80,4	10,3	3,1	6,2	0,0	97	8	12,38	8,07
3- lavelinge	73,8	7,5	7,9	10,7	0,0	240	17	13,41	5,04
4- école	81,9	6,2	2,8	8,4	0,7	144	19	7,58	4,60
5- manchots	75,6	6,1	0,0	15,8	2,4	82	8	9,75	6,50
6- pompier	76,5	9,4	9,5	4,7	0,0	85	7	11,57	5,32
7- camping	81,4	8,1	4,7	4,8	1,2	86	7	11,00	4,00
8- dinosaure	71,1	5,3	11,9	10,0	2,0	152	11	13,18	6,35
9- colibri	72,8	5,3	7,7	13,2	1,2	169	10	16,00	6,25
10- origine	87,1	3,4	4,3	5,2	0,0	116	7	16,14	4,60
Moyennes	**77,2**	**7,3**	**5,9**	**8,9**	**0,8**	**135,0**	**11,3**	**12,0**	**5,6**

Chaque texte est suivi de quatre énoncés sur lesquels les sujets doivent poser un jugement. On demande au sujet d'appuyer sur O, pour oui, quand l'énoncé « représente le texte que tu viens de lire », ou sur N dans le cas contraire. Les réponses ont été construites de façon à ce que les « meilleures » réponses ne soient pas systématiquement les plus longues ou les plus élaborées, afin de pallier la possibilité de réussir le test en sélectionnant les réponses avec des concepts plus abstraits ou des mots plus rares. Les modalités du pointage, correspondant à plusieurs qualités de réponses, représentent les trois « niveaux » de représentation ayant lieu pendant le processus de compréhension de texte (van Dijk et Kintsch, 1983).

- Trois points sont accordés aux réponses correspondant à la construction d'un 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. Ce modèle est une représentation mentale, ou structure, cohérente de la situation dénotée par le récit (Fayol, 2003; Gernsbacher, 1990, 1991; Kintsch et van Dijk, 1978; Oakhill et Garnham, 1988; van Dijk et Kintsch, 1983). Pour rappel, le modèle mental représente véritablement la situation décrite par le texte, plutôt que le texte lui-même.

- Deux points sont accordés à une réponse qui dénote de la construction de la base propositionnelle du texte (Kintsch et van Dijk, 1978; van Dijk et Kintsch, 1983) : des inférences sont faites à partir des éléments qui n'étaient pas explicites dans le texte. On dépasse ici la compréhension littérale du texte.

- Un point est accordé à une réponse vraie, qui dénote une compréhension superficielle, ou littérale, du récit. C'est la construction d'une structure de surface du texte (Kintsch et van Dijk, 1978; van Dijk et Kintsch, 1983), une compréhension de ce qui est explicite dans le texte, qui est mesurée : les éléments lexicaux et syntaxiques du texte sont reconnus. Ce niveau structurel de compréhension de texte correspond chez Giasson (1990) à la microstructure : les phrases du texte et leur relation entre elles sont reconnues. Une réponse est littérale si elle constitue un synonyme d'une information du texte, ou si elle présente des similarités sémantiques évidentes avec un élément du texte lu (Giasson, 1990).

Les réponses fausses valent zéro point. Celles-ci reprennent des éléments du récit, comme des personnages et des actions, en changeant les liens qui les unissent. Elles dénotent chez le sujet une incapacité à extraire le sens du texte. Ces réponses jouent aussi un rôle de leurres, et représentent une forme de pointage anti-hasard : en représentant jusqu'à deux réponses sur quatre dans les choix, elles limitent la sélection aléatoire des meilleures réponses.

Neuf juges expert·es en didactique du français ou en linguistique se sont penché·es sur la conception de l'épreuve, afin d'éprouver la validité de notre jugement quant au pointage de l'épreuve. Elles et ils ont été recruté·es à l'Université du Québec à Montréal et à la Haute École pédagogique du canton de Vaud. Nous leur avons demandé de lire les 10 textes de l'épreuve et d'attribuer une note de 0 à 3 points pour chacun des quatre énoncés proposés, en fonction de l'échelle suivante :

- 0 : l'énoncé ne se rapporte pas du tout au texte ou il reprend des éléments du récit, comme des personnages et des actions, en changeant les liens qui les unissent dans le récit. Il s'agit d'une réponse « fausse ».
- 1 : l'énoncé dénote une compréhension superficielle, ou littérale, du récit. Les phrases du texte et leur relation entre elles sont reconnues. L'énoncé est synonyme d'une information du texte ou présente des similarités sémantiques évidentes avec un élément du texte.
- 2 : l'énoncé se rapporte au texte et dénote que des inférences sont faites à partir des éléments qui n'étaient pas explicites dans le texte. Ils relèvent de l'implicite, ou d'inférences logiques faites par la lectrice ou le lecteur. L'énoncé se réfère à des éléments que le texte ne dit pas, mais qu'on déduit.
- 3 : l'énoncé se rapporte au texte et représente une compréhension du texte qui est le produit d'interactions entre la phrase et sa représentation, l'information textuelle, ainsi que la connaissance qu'a le lecteur du monde. Il s'agit d'une représentation cohérente de la situation dénotée par le récit. Il

s'agit de la « meilleure » réponse, ou celle qui va au-delà du texte, et qui dénote un niveau de compréhension supérieur.

Les juges ont été informé·es du fait que les quatre niveaux de réponses de l'échelle détaillée précédemment n'étaient pas nécessairement tous représentés dans les choix de réponses relatifs à chaque texte, et qu'une même « note » (correspondant à un niveau de réponse) pouvait être attribuée plus d'une fois dans le même bloc de réponses. Les juges avaient donc pour consigne de réfléchir à la « note » à donner à chaque énoncé, sans procéder par élimination.

Le calcul de fiabilité interjuges pour cette épreuve nous a fourni un alpha de Cronbach de 0,943, ce qui indique qu'elle a un niveau élevé de cohérence et de fiabilité : la corrélation interitems est élevée. Par ailleurs, cet alpha élevé indique aussi que l'échelle de mesure possède une excellente cohérence interne, et que l'outil semble mesurer un même construit théorique.

6.2.4.2 Description de la variabilité dans les textes

Bien que nous ayons montré que les textes avaient un haut niveau de cohérence sur les plans de la fréquence lexicale, de la longueur des phrases, de la longueur totale et du nombre de mots par phrase, éléments que nous souhaitons contrôler, nous avons voulu vérifier si les textes induisaient des taux de réussite différents. Cette variabilité dans l'échantillon de données est nécessaire à l'application d'analyses

inférentielles. Des ANOVA à mesures répétées ont été effectuées pour vérifier cette hypothèse. Les résultats montrent que les taux de réussite sont effectivement fonction des différents textes ($F(9, 1710) = 73.1$, $p < 0{,}001$, $\eta^2 = 0.222$).

Tableau 6.8 Statistiques descriptives pour les différents petits textes

	TXT1	TXT2	TXT3	TXT4	TXT5	TXT6	TXT7	TXT8	TXT9	TXT10	COMPpetitstextes
N	191	191	191	191	191	191	191	191	191	191	191
Manquants	18	18	18	18	18	18	18	18	18	18	18
Moyenne	0.662	0.871	0.716	0.529	0.572	0.757	0.639	0.286	0.455	0.657	0.654
Moyenne de l'erreur-standard	0.0242	0.0189	0.0207	0.0207	0.0219	0.0230	0.0225	0.0149	0.0201	0.0234	0.00854
Moyenne des IC à 95% limite inférieure	0.615	0.834	0.676	0.488	0.529	0.712	0.594	0.257	0.416	0.611	0.637
Moyenne des IC à 95% limite supérieure	0.710	0.908	0.757	0.570	0.615	0.802	0.683	0.315	0.495	0.703	0.671
Médiane	0.833	1.00	0.833	0.500	0.600	0.800	0.600	0.200	0.571	0.500	0.664
Écart-type	0.335	0.261	0.286	0.285	0.302	0.318	0.311	0.206	0.278	0.323	0.118
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.278
Maximum	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.800	1.00	1.00	0.940

Les moyennes de réussite pour les textes indiquent clairement un effet plafond pour le texte 2, et un effet plancher pour le texte 8, ce qui est corroboré par les tests post hocs (voir annexe K). Mises à part les différences marquées entre les textes 2, 8 et les autres textes, les tests post hocs indiquent également que les différences de moyennes entre les textes 4, 9 et la plupart des autres textes sont à la source de la significativité de l'anova. Les corrélations (*r* de Pearson bilatéral) entre les dix textes sont unilatéralement faibles, se situant entre 0,15 et 0,338, en excluant les corrélations non significatives (voir l'annexe K, où sont rapportées les analyses détaillées ayant trait aux textes courts).

Nous devons mentionner un biais expérimental évident, inhérent à l'idée de mesurer le modèle de situation par des choix de réponses. La lecture d'un texte, quel qu'il soit, induit autant de modèles de situation qu'il existe de lectrices et lecteurs. Les modèles de situation sont fonction des compétences langagières, de l'historique personnel, du milieu culturel et de tout ce qui constitue l'individu dans son individualité et sa spécificité (Farrall, 2012). La mesure d'un construit qui se caractérise par la diversité de ses manifestations est notoirement périlleuse.

L'ajout d'une épreuve sur texte long, aménagée avec des questions à choix dont les modalités de réponses sont moins fixes, concourt à la mesure d'une petite partie de la subjectivité propre à la construction du modèle de situation par l'individu. Cette épreuve fait l'objet de la section suivante.

6.2.4.3 Épreuve avec texte long

Le matériel expérimental élaboré pour cette épreuve de compréhension de textes longs correspond à un texte narratif, donc un texte qui raconte. Le texte a été traduit et adapté de l'une des quatre nouvelles incluses dans la *Multi-media Comprehension Battery* (Gernsbacher et Varner, 1988). Il s'agit d'une sorte de fable, puisqu'il contient une morale, qu'il s'agit d'interpréter. Le texte a, de plus, cette particularité de décrire un espace (une ferme) dans lequel les événements se déroulent. Sans une création assez précise d'un modèle mental, ou carte mentale, de cette maison dans sa situation spatiale (Bestgen, 2007), le sujet ne peut comprendre le caractère comique des événements ni les émotions et les motivations des personnages.

L'épreuve a été élaborée en parallèle de l'épreuve sur textes courts, et dans le but de la compléter, et d'offrir une mesure plus fiable des performances en compréhension chez les sujets. En effet, l'épreuve sur textes courts, avec questions à choix, est susceptible de mesurer autre chose que la compréhension en lecture, puisque son format à choix permet une bonne performance en ne comptant que sur le hasard (Farrall, 2012).

L'épreuve comprend un texte long, sur deux pages, et plusieurs réponses ouvertes ou libres, en plus d'autres modalités de réponses. Les tests avec réponses ouvertes ou libres présentent le désavantage d'avoir une fidélité réduite, par les problèmes de cotation qu'ils posent (plusieurs réponses peuvent s'avérer correctes) (Farrall, 2012).

Le texte fait deux pages, soit 878 mots (58 phrases, pour une moyenne de 14,93 mots par phrase) et est complété par 11 questions variées, qui figurent à l'annexe G. Les questions, originales, portent tour à tour sur des détails précis, sur les relations entre les personnages, l'ordre des événements de l'histoire, etc. Un tiers des réponses mesurent la reconnaissance des éléments lexicaux et syntaxiques explicitement présents dans le texte (« Comp 1 »). Un autre tiers mesure des éléments qui sont le résultat des inférences produites sur la base d'éléments qui n'étaient pas explicitement présents dans le texte (« Comp 2 »). Le tiers restant cherche à mesurer, chez les sujets, des informations générées sur la base du lien entre les deux premiers niveaux et les connaissances du monde, afin de générer un modèle mental cohérent (« Comp 3 »). Ces regroupements de questions sont inspirés par les évaluations en compréhension chez Foucambert (2003). Les 12e et 13e questions ne sont pas comptabilisées dans les résultats : elles ont été ajoutées afin de permettre aux sujets plus rapides de continuer à travailler sur l'épreuve sans déranger le reste du groupe.

All models are wrong, but some are useful.

— George Box

CHAPITRE VII

DESIGN EXPÉRIMENTAL

Ce chapitre est consacré à la présentation du design expérimental utilisé pour répondre à nos questions de recherche et complète la présentation des différents outils élaborés pour mesurer les construits qui nous intéressent. Seront décrits au sein de ce chapitre le design expérimental de notre étude (7.1), les prétests (7.2) et les traitements statistiques prévus pour l'analyse des données (7.3).

7.1 Design expérimental

Notre étude, de type transversal, vise à évaluer la contribution de variables mesurées et de variables latentes à la compréhension en lecture. Il s'agit plus précisément d'explorer la structure factorielle sous-jacente et l'envergure des relations entre les habiletés métalinguistiques et la compréhension en lecture, à l'instar notamment de Kieffer *et al.* (2016), Kim *et al.* (2019), Li et Wu (2015) et Tighe *et al.* (2018).

Notre devis expérimental comporte, en plus d'une phase d'analyses descriptives, deux séries d'analyses inférentielles, paramétriques et confirmatoires effectuées à l'aide de modèles travaillés par SEM (*Structural Equation Modeling*). Le traitement des données reposera donc sur deux moments consécutifs. Dans le premier, le modèle de mesure, la relation entre les variables latentes – c'est-à-dire les trois construits ainsi que leurs dimensions hypothétisées – et leurs indicateurs (données mesurées) sera démontré. Cette étape nous permettra de vérifier expérimentalement la validité de nos variables latentes. Dans le deuxième, les contributions des HmL (variables indépendantes) aux performances en compréhension en lecture (variables dépendantes) seront analysées, puis les contributions des variables latentes issues des différentes tâches mesurant les HmL (variables indépendantes) aux performances en compréhension en lecture (variables dépendantes) seront analysées. Ce sont ces deux dernières séries d'analyses qui nous permettront de répondre à nos deux premières questions de recherche.

Afin de vérifier l'impact des différences interindividuelles sur les résultats, les données socioéconomiques de chaque sujet sont colligées (âge, langue première). Tou·tes les élèves ayant été inclus·es dans l'étude sont normolectrices et normolecteurs. L'impact du bilinguisme sur les données n'est pas pris en compte : les élèves bilingues n'ont pas été exclu·es d'office des données.

7.2 Prétests

Tous les instruments de mesure ont été prétestés en novembre et décembre 2019 préalablement à la collecte de données. Ceux-ci ont été présentés aux élèves en format papier, étant donné que seuls les items finalisés devaient être inclus dans l'élaboration des épreuves sur ordinateur. Les modifications ayant été apportées aux épreuves sont les suivantes :

- Recours à des mots plus fréquents dans les deux épreuves de lecture et leurs questions de compréhension associées.
- Séparation des épreuves de réplication de l'erreur en fonction de l'opération mobilisée. Cette modification s'est avérée nécessaire, d'une part, parce que la diversification des stratégies de réplication représentait une réelle difficulté pour les élèves (n'étant à priori pas liée à leurs habiletés en conscience syntaxique, mais bien à leur compréhension de la consigne), et d'autre part parce qu'il était beaucoup plus aisé d'implémenter séparément les trois types d'épreuves dans le logiciel permettant de les construire et de les lancer dans *Psychopy* (Peirce *et al.*, 2019).

- Uniformisation des items de conscience syntaxique afin qu'ils soient conformes aux consignes du logiciel (ex. déplacement des constituants les plus petits, sauf dans le cas des subordonnées relatives, etc.)
- Clarification des consignes au sein du logiciel.

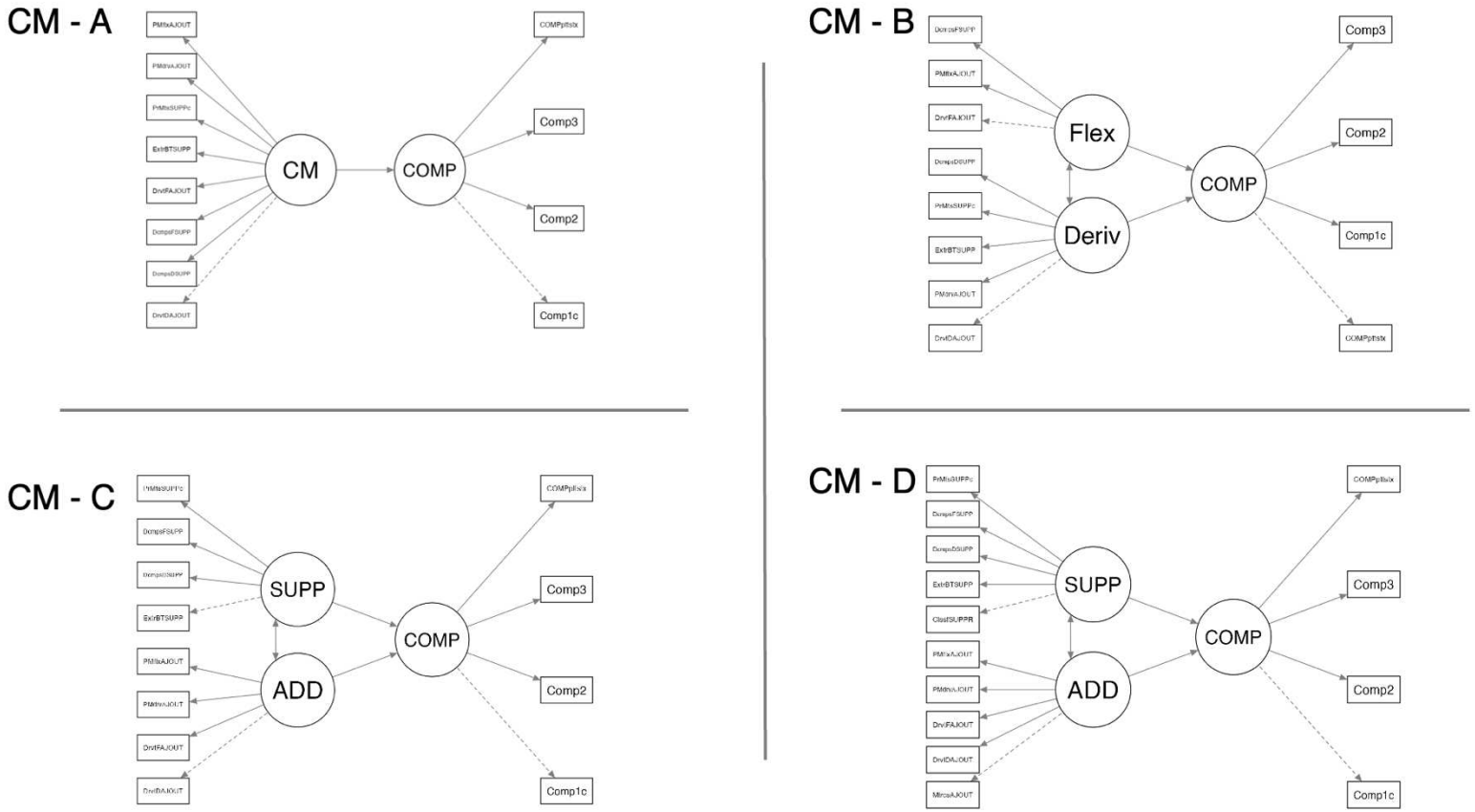
7.3 Traitement des données

Le projet cherche à évaluer heuristiquement des variables latentes représentant des compétences transversales aux épreuves. Les analyses descriptives d'usage seront réalisées et présentées : taux de réussite aux épreuves, normalité des distributions de chacune des variables et analyses de corrélation. À l'instar des études récentes portant sur le rapport entre habiletés psycholinguistiques et lecture (Anthony et Lonigan, 2004; Simard, D. *et al.*, 2012b), le projet présentera un modèle statistique travaillé par SEM à l'aide des logiciels Jamovi (The jamovi project, 2021) et R (R Core Team, 2020), faisant tous les deux fonctionner le paquet *lavaan* (Rosseel, 2012). Ces traitements statistiques permettront d'évaluer les modèles que nous construisons à partir de la littérature scientifique et de les mettre en relation avec la performance générale en lecture.

Les figures 7.1 et 7.2 constituent des représentations graphiques de deux types de dimensions à évaluer à l'aide des modélisations par SEM : des variables latentes regroupant des unités linguistiques minimales, d'une part, et des variables latentes regroupant des opérations cognitives minimales d'ajout et de suppression, d'autre

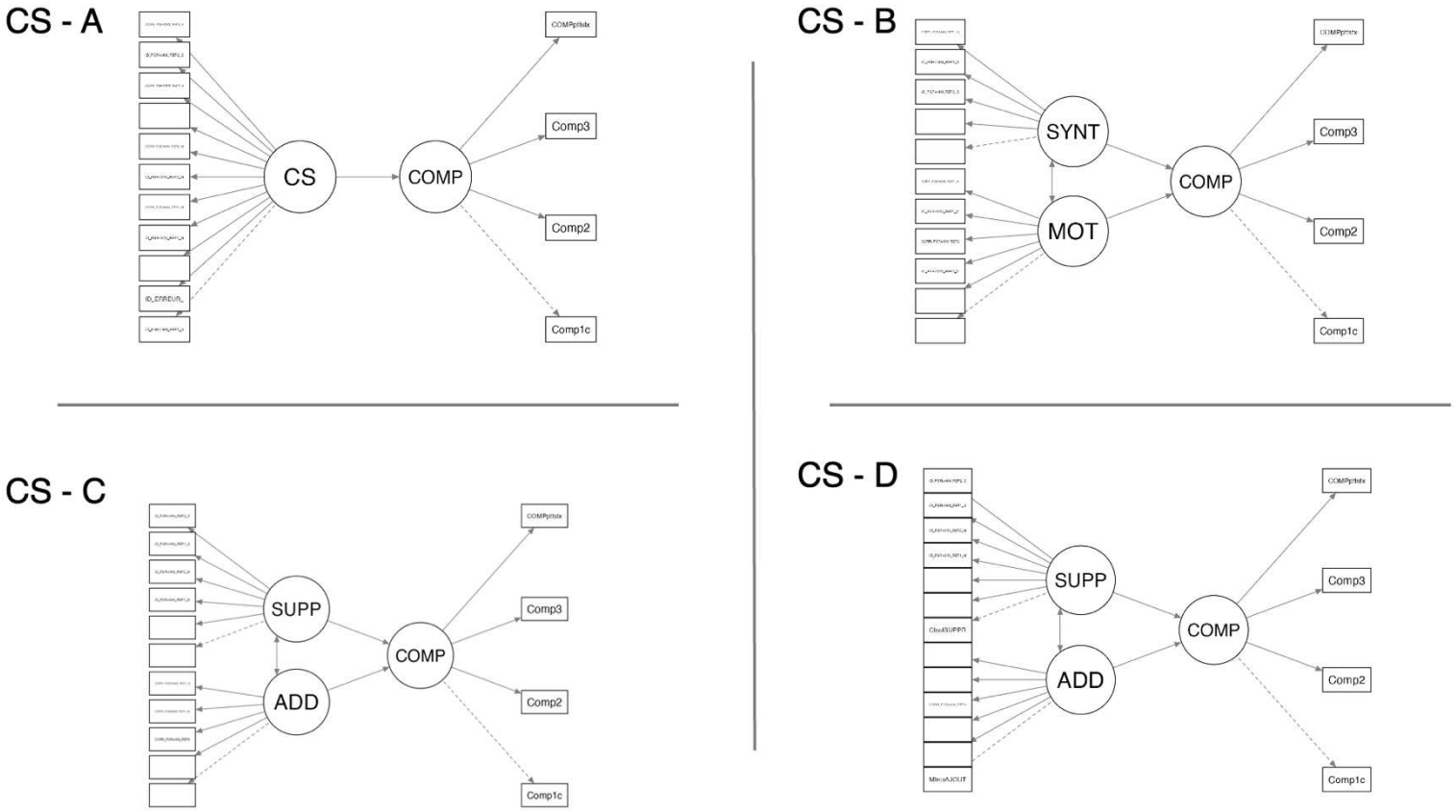
part. Les modèles unidimensionnels relatifs aux deux construits et les modèles regroupant les opérations cognitives avec, en plus, les variables matricielles, seront également évalués. Ainsi, quatre modèles contrastés sont testés en regard des autres, et ce pour les deux construits :

- Modèle A : modèle unidimensionnel ;
- Modèle B : modèle bidimensionnel impliquant deux types d'unités linguistiques ;
- Modèle C : modèle bidimensionnel impliquant deux types d'opérations cognitives ;
- Modèle D : modèle bidimensionnel impliquant à nouveau deux types d'opérations cognitives, mais avec les deux épreuves matricielles non-linguistiques.



Note. CM-A : modèle unidimensionnel ; CM-B : modèle bidimensionnel Unités ; CM-C : modèle bidimensionnel Opérations ; modèle CM-D : modèle bidimensionnel Opérations avec épreuves matricielles

Figure 7.1 Représentations graphiques des modèles à tester, en fonction des dimensions hypothétisées dans les épreuves de conscience morphologique



Note. CM-A : modèle unidimensionnel ; CM-B : modèle bidimensionnel Unités ; CM-C : modèle bidimensionnel Opérations ; modèle CM-D : modèle bidimensionnel Opérations avec épreuves matricielles

Figure 7.2 Représentations graphiques des modèles à tester, en fonction des dimensions hypothétisées dans les épreuves de conscience syntaxique

PARTIE 3 RÉSULTATS ET PERSPECTIVES

[…] truth is only to be had by laying together many varieties of error.

— Virginia Woolf, *A Room of One's Own*

CHAPITRE VIII

ANALYSES DESCRIPTIVES ET INFÉRENTIELLES DES DONNÉES

Ce chapitre sera consacré aux analyses des données recueillies afin de mettre au jour les différences individuelles pour les trois construits étudiés, soit la conscience morphologique, la conscience syntaxique et la compréhension en lecture. Ces différences individuelles seront vues sous le prisme des opérations de suppression et d'addition ayant présidé à la constitution des différentes épreuves de conscience linguistique. Après avoir rappelé nos questions de recherches (8.1), nous ferons état de la manière dont nous avons préparé les données (8.2) préalablement aux différentes analyses, effectuées en deux temps. Dans un premier temps seront présentés dans ce chapitre les résultats des statistiques descriptives (8.3), des corrélations (8.4) et des analyses en composantes principales (8.5). Dans un deuxième temps, nous procéderons à la description des spécifications des modèles SEM (8.6) et du plan d'analyses (8.7) pour la conscience morphologique (8.8) et la conscience syntaxique (8.9). En 8.10, nous ferons un résumé des résultats et des limites de la recherche.

8.1 Rappel des questions de recherche

Pour rappel, nos questions de recherche de nature psycholinguistique, qui guideront les analyses subséquentes, sont les suivantes :

- *Les mesures traditionnelles de la conscience morphologique et de la conscience syntaxique peuvent-elles se scinder en deux constituants, prenant la forme de variables latentes ?*

- *Dans quelle mesure les variables latentes construites – d'ordres linguistique et cognitif – contribuent-elles à la compréhension en lecture?*

8.2 Préparation des données

Chaque sujet s'est vu accorder un code alphanumérique afin d'assurer l'anonymat des données.

8.2.1 Préparation des fichiers de données

Les données, dans leur état le plus brut, sont organisées de façon hiérarchique : chaque élève possède un résultat pour chacun des items de chacune des épreuves (plusieurs lignes par sujet). Ce résultat est binaire pour les 11 épreuves mesurant la conscience morphologique et la conscience syntaxique, permettant notamment des

analyses logistiques, mais il se déploie de manière plus nuancée pour les deux épreuves de compréhension[71]. À défaut d'une méthode plus sophistiquée, les données manquantes au sein de chacune des épreuves ont été traitées comme des erreurs (*zero imputation method*) (Dai, 2021), avant que les variables ne soient ramenées à des moyennes de réussite par sujet. Nous avons ainsi tenu pour acquis qu'un sujet laissant une question sans réponse n'était en fait pas en mesure d'y répondre, ou pas en mesure d'y répondre assez rapidement (Dai, 2021).

Ensuite, pour faciliter le traitement des données en vue des analyses de corrélation et des SEM, les résultats pour chaque épreuve et chaque sujet ont tous été codés avec un ratio de 1, avec une ligne de résultats par sujet, afin d'obtenir des variables continues comparables. Celles-ci correspondent aux résultats des sujets aux épreuves matricielles avec matériau non linguistique, aux épreuves de conscience morphologique et aux épreuves de conscience syntaxique – divisées en fonction des unités et des opérations qu'elles mettent en œuvre – et des deux épreuves de compréhension en lecture, pour un total de 37 variables.

Par la suite, nous avons vérifié la normalité des distributions de ces variables. La normalisation des données constitue un préalable à l'utilisation de certaines des statistiques paramétriques auxquelles les données doivent se prêter, notamment les corrélations et les régressions (Field *et al.*, 2012; Hahs-Vaughn et Lomax, 2020). Nous

[71] Voir la section 6.2.4 pour les informations docimologiques à propos des deux épreuves de compréhension.

avons procédé à des analyses descriptives de nos données afin de décrire les distributions de nos variables et de vérifier leur normalité. Ces analyses descriptives figurent dans le tableau 8.1. Les valeurs d'asymétrie et d'aplatissement sont toutes bien à l'intérieur de la fourchette des +-3 pour les SEM recommandés par Xuelian *et al.* (2019), ou même des valeurs d'asymétrie et d'aplatissement entre +-2 pour Hahs-Vaughn et Lomax (2020), indiquant la présence de variables plutôt normalement distribuées. La forme des distributions est toutefois plus pertinente pour déterminer leur normalité dans un échantillon de 200 ou plus (Field *et al.*, 2012). L'inspection des graphiques montre plusieurs distributions avec asymétrie négative (la majorité des données se retrouvent à droite de la courbe), en particulier pour ce qui est des variables mesurant la conscience morphologique. Les hypothèses de normalité étant enfreintes, nous avons transformé les variables avec des valeurs d'asymétrie de +-1 et plus en les mettant au carré, puis avons adapté les méthodes d'estimation dans la suite de nos analyses (Loehlin, 2003). Les valeurs pour les nouvelles variables créées figurent dans le tableau 8.2.

8.2.2 Procédure

Les données ont été récoltées dans les écoles entre novembre 2019 et décembre 2020, lors de sessions collectives d'exercices à l'ordinateur, tenues durant les heures régulières de classe. À l'exception de rares conflits d'horaires, les 5 ou 6 séances de collecte de données (en fonction de la rapidité des élèves) se sont déroulées à la même heure, semaine après semaine, pour chacune des 12 classes. La pandémie et

le confinement qui s'en est suivi en 2020 ont retardé la collecte des données, sans pour autant occasionner de biais expérimental majeur ; seules les séances de passation d'une des classes de 5P ont été interrompues par le confinement. Les autres passations ont pu avoir lieu avant ou après ces évènements.

8.2.3 Description de l'échantillon

Nous avons récolté les données sociodémographiques suivantes au sein de notre échantillon d'élèves : nom, prénom, date de naissance, langue parlée à la maison, langue première de la mère, langue première du père[72]. Tou·tes les élèves font leur scolarité dans des écoles du canton de Vaud, en Suisse. Plusieurs parents d'élèves n'ont pas souhaité fournir les informations demandées, et nous n'avons pu obtenir de renseignements sur l'âge des élèves pour deux classes. Nous avons donc estimé l'âge des élèves en fonction de leur degré scolaire.

L'échantillon, après l'exclusion de 4 élèves dont les parents n'ont pas autorisé leur participation à l'étude, et de 5 autres, exclus pour des troubles d'apprentissage diagnostiqués, comprend 208 élèves (91 filles : 43,8 % du total, 117 garçons : 56,3 % du total), scolarisés dans 12 classes parmi 5 écoles du canton de Vaud, en Suisse

[72] Voir le formulaire de consentement type remis aux enseignant·es à l'annexe I. [73] L'exclusion des valeurs qui sont plus de trois fois l'écart type de la moyenne est une pratique couramment utilisée pour l'identification et le traitement des données aberrantes en statistiques, et est appelée la « règle des 3-sigma ». L'utilisation de cette règle repose sur les valeurs de la distribution normale canonique, dans laquelle environ 68 % des données s'écartent d'un écart type de la moyenne, 95 % se situent à l'intérieur de deux écarts types de la moyenne, et 99 % sont sous trois écarts types de la moyenne (Hahs-Vaughn et Lomax, 2020, p.1025).

romande. L'âge moyen des élèves est de 10,3 ans (min = 8 ; max = 12,6 ; é-t = 1,25). Le choix de cette tranche d'âge est guidé par le souci d'obtenir une variance suffisante dans les résultats : d'un côté en raison de la maitrise raisonnable de la langue écrite que possèdent les individus de cet âge et de l'autre, pour éviter les effets plafonds associés aux populations plus âgées (Gombert, 1990). Il est à noter que les données sur l'âge de 30 élèves ne sont pas disponibles. Les élèves se classent en quatre niveaux scolaires : 57 sont en 5P (27,4%), 63 sont en 6P (30,3 %), 58 sont en 7P (27,9 %), et 30 sont en 8P (14,4 %).

Notre échantillon fait état de nombreuses données manquantes en ce qui a trait au profil linguistique des élèves. En effet, les parents de 102 élèves, soit près de la moitié de l'échantillon, n'ont pas souhaité fournir d'informations sur la langue première de leur enfant. Chez les élèves dont nous avons obtenu des informations linguistiques (n=106), 87 parlent le français à la maison (83,7 %), 87 ont le français comme langue première, et 19 ont une autre langue première. Parmi les langues autres que le français parlées à la maison, on retrouve le portugais (n=4 ; 3,8 % du total), l'albanais (n=2 ; 1,9 % du total), l'italien (n=2 ; 1,9% du total), le kurde (n=2 ; 1,9 % du total). Le caractère lacunaire de ces données, couplé à un manque d'informations sur la maitrise effective du français chez les sujets, nous a convaincue d'analyser l'ensemble du corpus, sans séparer les francophones des allophones. Ce parti pris pour une étude des liens entre conscience grammaticale et compréhension sans égard à la

langue première des élèves est motivé par le désir de représenter fidèlement la réalité sociodémographique et linguistique des classes, en Suisse et ailleurs.

8.3 Statistiques descriptives pour les variables mesurées

Afin d'obtenir une vision globale de notre base de données et de détecter toute donnée aberrante éventuelle, nous avons réalisé des analyses descriptives approfondies de nos données. La liste des variables et de leurs dénominations figure à l'annexe M. Les temps de réponse aberrants pour l'ensemble des épreuves ont été traités de la même manière : les temps 3 fois au-dessus de la moyenne ont été exclus (Hahs-Vaughn et Lomax, 2020), représentant chaque fois moins de 1,5% d'exclusions de données[73]. Les élèves de notre échantillon lisent en moyenne un texte par minute – 57,867 secondes pour des textes qui font en moyenne 135 mots, pour une vitesse de lecture de 139,8 mots par minute.

Le tableau 8.1 indique les valeurs des 37 variables utilisées, après exclusion des données aberrantes. Les moyennes (taux de réussite) sont assez élevées pour les épreuves de conscience morphologique (entre 59,2% et 85,1%) et les épreuves logiques de classifications ($\bar{x}$ = 78,7%) et de matrices ($\bar{x}$ = 78,2%). L'épreuve de

[73] L'exclusion des valeurs qui sont plus de trois fois l'écart type de la moyenne est une pratique couramment utilisée pour l'identification et le traitement des données aberrantes en statistiques, et est appelée la « règle des 3-sigma ». L'utilisation de cette règle repose sur les valeurs de la distribution normale canonique, dans laquelle environ 68 % des données s'écartent d'un écart type de la moyenne, 95 % se situent à l'intérieur de deux écarts types de la moyenne, et 99 % sont sous trois écarts types de la moyenne (Hahs-Vaughn et Lomax, 2020, p.1025).

compréhension avec textes courts ($\bar{x}$ = 65,3 %) est mieux réussie que celle portant sur un texte long ($\bar{x}$ = 57,5 %). Une ANOVA à mesures répétées a permis de confirmer que les trois types de questions de compréhension sur texte long, correspondant à trois « niveaux » de compréhension (Comp 1, Comp 2 et Comp 3), ne sont pas équivalents (F (2,414) = 199,27, $p < 0{,}001$). Étonnamment, la moyenne de réussite pour l'item Comp 2 est plutôt bas ($\bar{x}$: 30,6 %), tandis que la moyenne relative au niveau réputé plus bas de compréhension (Comp 1, $\bar{x}$: 70,2 %) est plus proche du niveau réputé le plus sophistiqué (Comp 3, $\bar{x}$: 62,3 %), même si elles sont significativement différentes. Les épreuves de conscience syntaxique, quant à elles, sont les moins bien réussies ($\bar{x}$: 28,5 %-74,6 %). Conformément à ce qui est observé dans la littérature (Gaux et Gombert, 1999b), ce sont les variables regroupant les items de réplication de l'erreur qui affichent les résultats les plus faibles ($\bar{x}$ avec une première phrase grammaticale = 49,7 %, $\bar{x}$ avec une deuxième phrase grammaticale plus complexe = 44,6 %). Les mesures de correction de l'erreur ($\bar{x}$ = 65,5 %) affichent des moyennes de réussite légèrement inférieures à celles d'identification de l'erreur ($\bar{x}$: 68,6%).

Tableau 8.1 Statistiques descriptives des variables brutes

	N	Man-quants	Moy.	Méd.	Écart type	Min.	Max.	Coefficient d'asymétrie		Kurtosis	
								Asymétrie	Erreur standard	Kurtosis	Erreur standard
VARIABLES MESURANT LA COMPRÉHENSION											
Comp1	208	0	0.702	0.778	0.246	0.0000	1.000	-1.111	0.169	0.5244	0.337
Comp2	208	0	0.306	0	0.220	0.0000	1.000	0.6804	0.169	0.2264	0.337
Comp3	208	0	0.623	0.667	0.203	0.0000	1.000	-0.884	0.169	0.4664	0.337
moyenne T lecture	190	18	57.867	54.563	19.94	14.626	120.646	0.7072	0.176	0.2805	0.351
COMPpetitstextes	190	18	0.653	0.664	0.118	0.2778	0.940	-0.3464	0.176	0.2426	0.351
VARIABLES MATRICIELLES NON LINGUISTIQUES											
ClassificationSUPPR	185	23	0.787	0.800	0.156	0.2667	1.000	-1.1682	0.179	1.1933	0.355
MatricesAJOUT	185	23	0.782	0.800	0.178	0.0667	1.000	-1.1066	0.179	1.0854	0.355
VARIABLES MESURANT LA CONSCIENCE MORPHOLOGIQUE											
DecompositionDSUPP	200	8	0.814	0.875	0.194	0.0625	1.000	-1.5548	0.172	2.3574	0.342
DecompositionFSUPP	200	8	0.742	0.875	0.224	0.0000	1.000	-1.0930	0.172	0.5588	0.342
DecompositionTotalSUPP	200	8	0.790	0.833	0.192	0.0417	1.000	-1.4728	0.172	1.9507	0.342
DerivationDAJOUT	201	7	0.703	0.813	0.282	0.0000	1.000	-1.0874	0.172	0.2072	0.341
DerivationFAJOUT	201	7	0.785	0.875	0.244	0.0000	1.000	-1.4089	0.172	1.4757	0.341
DerivationTotalAJOUT	201	7	0.730	0.833	0.260	0.0000	1.000	-1.1942	0.172	0.5290	0.341
ExtractionBasePrefSUPP	200	8	0.765	0.833	0.234	0.0000	1.000	-1.2801	0.172	1.1363	0.342
ExtractionBaseSuffSUPP	200	8	0.729	0.750	0.232	0.0833	1.000	-0.7000	0.172	-0.390	0.342

	N	Man-quants	Moy.	Méd.	Écart type	Min.	Max.	Coefficient d'asymétrie		Kurtosis	
								Asymétrie	Erreur standard	Kurtosis	Erreur standard
ExtractionBaseTotalSUPP	200	8	0.747	0.792	0.207	0.0417	1.000	-1.0774	0.172	0.8406	0.342
PaireMotsSUPP	201	7	0.851	0.917	0.143	0.3333	1.000	-1.3294	0.172	1.1356	0.341
PMderivAJOUT	199	9	0.592	0.600	0.203	0.0000	1.000	-0.3402	0.172	-0.4526	0.343
PMflexAJOUT	202	6	0.671	0.688	0.218	0.0625	1.000	-0.6233	0.171	-0.2084	0.341
ChoixPMTotalAJOUT	199	9	0.632	0.646	0.191	0.1313	0.967	-0.4262	0.172	-0.5969	0.343
VARIABLES MESURANT LA CONSCIENCE SYNTAXIQUE											
ID_ERREUR_PAGRAMM_M	182	26	0.744	0.778	0.183	0.0000	1.000	-1.5201	0.180	2.4668	0.358
CORR_ERREUR_PAGRAMM_M	182	26	0.562	0.556	0.220	0.0000	1.000	-0.5220	0.180	-0.3236	0.358
ID_PGRAMM_REP1_M	182	26	0.518	0.556	0.260	0.0000	1.000	-0.0114	0.180	-0.8604	0.358
CORR_PGRAMM_REP1_M	185	23	0.680	0.667	0.266	0.0000	1.000	-0.5976	0.179	-0.4802	0.355
ID_PGRAMM_REP2_M	182	26	0.473	0.444	0.249	0.0000	1.000	0.2129	0.180	-0.6596	0.358
CORR_PGRAMM_REP2_M	185	23	0.668	0.667	0.292	0.0000	1.000	-0.5544	0.179	-0.6205	0.355
ID_ERREUR_PAGRAMM_S	184	24	0.627	0.667	0.187	0.0000	0.889	-1.0070	0.179	0.6208	0.356
CORR_ERREUR_PAGRAMM_S	184	24	0.746	0.778	0.240	0.0000	1.000	-1.1979	0.179	0.7924	0.356
ID_PGRAMM_REP1_S	184	24	0.432	0.444	0.250	0.0000	0.889	-0.0308	0.179	-0.9911	0.356

	N	Man-quants	Moy.	Méd.	Écart type	Min.	Max.	Coefficient d'asymétrie		Kurtosis	
								Asymétrie	Erreur standard	Kurtosis	Erreur standard
CORR_PGRAMM_REP1_S	192	16	0.359	0.333	0.323	0.0000	1.000	0.3978	0.175	-1.1634	0.349
ID_PGRAMM_REP2_S	184	24	0.365	0.333	0.263	0.0000	1.000	0.3594	0.179	-0.8470	0.356
CORR_PGRAMM_REP2_S	192	16	0.285	0.167	0.311	0.0000	1.000	0.8050	0.175	-0.6091	0.349
TOTAL_ADD_SUPPR	191	17	0.536	0.542	0.180	0.0417	1.000	0.1108	0.176	-0.1181	0.350
TOTAL_DEPLACEMENT	192	16	0.456	0.438	0.241	0.0000	0.958	0.1478	0.175	-0.9944	0.349
TOTAL_REMPLACEMENT	191	17	0.704	0.792	0.243	0.0000	1.000	-0.7002	0.176	-0.4260	0.350
IDENTIFICATION_ERREUR	182	26	0.686	0.722	0.165	0.0556	0.944	-1.3258	0.180	1.7064	0.358
CORRECTION_ERREUR	182	26	0.655	0.722	0.212	0.0000	1.000	-1.0390	0.180	0.5235	0.358
REPLICATION1_ERREUR	182	26	0.497	0.467	0.237	0.0000	0.967	0.0830	0.180	-1.0162	0.358
REPLICATION2_ERREUR	182	26	0.446	0.433	0.239	0.0000	1.000	0.2988	0.180	-0.8532	0.358

Tableau 8.2 Statistiques descriptives des variables ayant subi une transformation

	N	Man-quants	Moy.	Méd.	Écart type	Min.	Max.	Coefficient d'asymétrie		Kurtosis	
								Asymétrie	Erreur standard	Kurtosis	Erreur standard
Comp1c	208	0	0.553	0.605	0.287	0.00000	1.00	-0.360	0.169	-0.7895	0.336
ClassificationSUPPRc	185	23	0.644	0.640	0.220	0.07111	1.000	-0.610	0.179	-0.1951	0.355
MatricesAJOUTc	185	23	0.642	0.640	0.247	0.00444	1.000	-0.506	0.179	-0.6154	0.355
DecompositionDSUPPc	200	8	0.699	0.766	0.263	0.00391	1.000	-0.848	0.172	-0.1158	0.342
DecompositionFSUPPc	200	8	0.600	0.766	0.284	0.00000	1.000	-0.450	0.172	-0.7718	0.342
DecompositionTotalSUPPc	200	8	0.660	0.694	0.254	0.00174	1.000	-0.811	0.172	-0.2403	0.342
DerivationDAJOUTc	201	7	0.573	0.660	0.320	0.00000	1.000	-0.398	0.172	-1.0588	0.341
DerivationFAJOUTc	201	7	0.676	0.766	0.309	0.00000	1.000	-0.700	0.172	-0.6781	0.341
DerivationTotalAJOUTc	201	7	0.601	0.694	0.308	0.00000	1.000	-0.558	0.172	-0.9212	0.341
ExtractionBasePrefSUPPc	200	8	0.639	0.694	0.295	0.00000	1.000	-0.596	0.172	-0.8169	0.342
ExtractionBaseTotalSUPPc	200	8	0.601	0.627	0.269	0.00174	1.000	-0.387	0.172	-0.8202	0.342
PaireMotsSUPPc	201	7	0.745	0.840	0.217	0.11111	1.000	-0.974	0.172	0.0560	0.341
ID_ERREUR_PAGRAMM_Mc	182	26	0.586	0.605	0.227	0.00000	1.000	-0.742	0.180	-0.2901	0.358
CORR_ERREUR_PAGRAMM_Sc	184	24	0.613	0.605	0.297	0.00000	1.000	-0.495	0.179	-0.7486	0.356
IDENTIFICATION_ERREURc	182	26	0.498	0.522	0.195	0.00309	0.892	-0.606	0.180	-0.2293	0.358

8.4 Corrélations

La matrice de corrélation de Pearson *r* figurant dans le tableau 8.3 indique que toutes les corrélations entre les variables mesurant la conscience morphologique, la conscience syntaxique, les mesures non linguistiques et la compréhension en lecture sont positives et significatives au seuil de $p < 0{,}001$, à une exception près : la corrélation entre COMP1c et ClassificationSUPPRc, faible ($r = 0{,}19$) et significative au seuil de $p < 0{,}05$. Les variables matricielles non linguistiques sont modérément corrélées ($r = 0{,}52$; $p < 0{,}001$). Les épreuves de compréhension en lecture présentent des coefficients de corrélation plutôt modérés également ($r = 0{,}29$-$0{,}56$; $p < 0{,}001$), marquant le caractère multifacette du construit mesuré. Notons que la variable COMP1c, qui représente une compréhension superficielle du texte long, présente une corrélation moins élevée avec les épreuves de CM et de CS que les variables mesurant les niveaux plus sophistiqués de la compréhension (Comp2, Comp3 et COMPpetitstextes).

Du côté des variables mesurant les HmL, les corrélations sont en général modérées à fortes. Les épreuves de conscience morphologique présentent des coefficients de corrélation de 0,51 à 0,81 ($p < 0{,}001$). L'empan des coefficients de corrélation relatifs aux épreuves de conscience syntaxique est assez large, allant de 0,45 à 0,90 ($p < 0{,}001$). Pour les deux construits métalinguistiques, souvent, les liens les plus forts peuvent être observés entre les variables partageant les mêmes formats, par exemple

le choix de suffixes dans des pseudomots dans les épreuves PMderivAJOUT et PMflexAJOUT (r = 0,65 ; p <0,001), les épreuves de complétion de phrases (DécompositionDSUPPc, DécompositionFSUPPc, DérivationDAJOUTc et DérivationFAJOUTc : r = 0,72-0,81 ; p <0,001) ou encore les segments dans lesquels les erreurs doivent être reproduites dans les épreuves de réplication d'erreurs (CORR_PGRAMM_REP1_M et CORR_PGRAMM_REP2_M r = 0,83 ; p <0,001 ; CORR_PGRAMM_REP1_S et CORR_PGRAMM_REP2_S : r = 0,90 ; p <0,001).

Pour rappel, une liste des épreuves et de leur dénomination au sein du présent chapitre se trouve à l'annexe M.

Tableau 8.3 Corrélations entre les variables à l'étude

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Comp1c	—												
2 Comp2	0.46	—											
3 Comp3	0.53	0.56	—										
4 COMPpetitstextes	0.28	0.38	0.32	—									
5 ClassificationSUPPRc	0.19	0.37	0.39	0.29	—								
6 MatricesAJOUTc	0.30	0.39	0.41	0.35	0.52	—							
7 DecompositionDSUPPc	0.33	0.45	0.45	0.43	0.43	0.37	—						
8 DecompositionFSUPPc	0.36	0.45	0.43	0.36	0.38	0.33	0.72	—					
9 DerivationDAJOUTc	0.36	0.46	0.47	0.48	0.40	0.35	0.76	0.70	—				
10 DerivationFAJOUTc	0.36	0.42	0.49	0.43	0.36	0.27	0.73	0.70	0.81	—			
11 ExtractionBaseTotalSUPPc	0.34	0.45	0.44	0.44	0.35	0.33	0.61	0.56	0.66	0.64	—		
12 PaireMotsSUPPc	0.50	0.41	0.48	0.52	0.32	0.43	0.57	0.51	0.60	0.53	0.52	—	
13 PMderivAJOUT	0.28	0.43	0.42	0.41	0.41	0.38	0.61	0.53	0.56	0.53	0.56	0.51	—
14 PMflexAJOUT	0.36	0.43	0.47	0.39	0.43	0.41	0.65	0.61	0.61	0.57	0.57	0.58	0.65
15 ID_ERREUR_PAGRAMM_Mc	0.36	0.37	0.50	0.46	0.34	0.38	0.45	0.46	0.53	0.50	0.45	0.52	0.44
16 CORR_ERREUR_PAGRAMM_M	0.35	0.39	0.48	0.55	0.43	0.47	0.53	0.44	0.55	0.53	0.52	0.60	0.54
17 ID_PGRAMM_REP1_M	0.33	0.46	0.46	0.45	0.43	0.43	0.54	0.50	0.57	0.53	0.51	0.55	0.50
18 CORR_PGRAMM_REP1_M	0.27	0.43	0.39	0.41	0.36	0.37	0.47	0.51	0.57	0.46	0.50	0.49	0.51
19 ID_PGRAMM_REP2_M	0.24	0.48	0.43	0.42	0.46	0.41	0.50	0.53	0.59	0.53	0.57	0.50	0.53
20 CORR_PGRAMM_REP2_M	0.28	0.39	0.38	0.38	0.39	0.31	0.43	0.47	0.54	0.44	0.48	0.46	0.47
21 ID_ERREUR_PAGRAMM_S	0.39	0.41	0.47	0.42	0.36	0.38	0.50	0.46	0.55	0.52	0.50	0.56	0.54
22 CORR_ERREUR_PAGRAMM_Sc	0.36	0.40	0.42	0.48	0.37	0.35	0.56	0.49	0.54	0.53	0.49	0.56	0.53
23 ID_PGRAMM_REP1_S	0.37	0.49	0.47	0.41	0.42	0.46	0.47	0.46	0.57	0.51	0.51	0.49	0.44
24 CORR_PGRAMM_REP1_S	0.33	0.45	0.38	0.40	0.39	0.39	0.44	0.40	0.52	0.44	0.45	0.44	0.39
25 ID_PGRAMM_REP2_S	0.35	0.44	0.44	0.41	0.42	0.46	0.43	0.48	0.52	0.47	0.49	0.42	0.47
26 CORR_PGRAMM_REP2_S	0.30	0.43	0.34	0.38	0.40	0.40	0.38	0.38	0.45	0.39	0.40	0.39	0.40

	14	15	16	17	18	19	20	21	22	23	24	25	26
1 Comp1c													
2 Comp2													
3 Comp3													
4 COMPpetitstextes													
5 ClassificationSUPPRc													
6 MatricesAJOUTc													
7 DecompositionDSUPPc													
8 DecompositionFSUPPc													
9 DerivationDAJOUTc													
10 DerivationFAJOUTc													
11 ExtractionBaseTotalSUPPc													
12 PaireMotsSUPPc													
13 PMderivAJOUT													
14 PMflexAJOUT	—												
15 ID_ERREUR_PAGRAMM_Mc	0.58	—											
16 CORR_ERREUR_PAGRAMM_M	0.60	0.80	—										
17 ID_PGRAMM_REP1_M	0.57	0.60	0.67	—									
18 CORR_PGRAMM_REP1_M	0.58	0.63	0.65	0.81	—								
19 ID_PGRAMM_REP2_M	0.61	0.65	0.66	0.81	0.74	—							
20 CORR_PGRAMM_REP2_M	0.56	0.62	0.60	0.73	0.83	0.79	—						
21 ID_ERREUR_PAGRAMM_S	0.59	0.60	0.64	0.51	0.54	0.49	0.54	—					
22 CORR_ERREUR_PAGRAMM_Sc	0.60	0.60	0.67	0.59	0.59	0.53	0.54	0.88	—				
23 ID_PGRAMM_REP1_S	0.58	0.58	0.62	0.63	0.59	0.68	0.61	0.66	0.67	—			
24 CORR_PGRAMM_REP1_S	0.50	0.55	0.59	0.65	0.58	0.65	0.60	0.65	0.68	0.87	—		
25 ID_PGRAMM_REP2_S	0.53	0.57	0.59	0.61	0.58	0.65	0.58	0.62	0.65	0.88	0.81	—	
26 CORR_PGRAMM_REP2_S	0.45	0.49	0.52	0.62	0.54	0.61	0.55	0.57	0.63	0.79	0.90	0.86	—

Note. Toutes les corrélations sont significatives avec un seuil de p<0,001 (sauf *r* COMP1c et ClassificationSUPPRc, significative au seuil de p<0,05).

8.5 Analyses en composantes principales

Des analyses en composantes principales avec rotation Varimax, de nature exploratoire, ont été effectuées afin de voir si des dimensions émergaient (Kunnan, 1998), sans hypothèses préconçues, parmi les mesures transformationnelles avec matériau non linguistique, de conscience morphologique et de conscience syntaxique. Ces analyses, comme toutes celles impliquant des variables latentes[74], présentent l'avantage de permettre une réduction des données (*data reduction*) et de leur complexité interprétative (Berger, J.-L., 2021).

Dans un premier temps, nous avons soumis les données à une ACP avec rotation varimax, avec la méthode d'extraction d'un nombre de composantes traditionnellement utilisée, soit celle avec un seuil de valeurs propres supérieur à 1 (Berger, J.-L., 2021; Field *et al.*, 2012). Dans ces conditions, quatre types d'épreuves émergent : les épreuves de CM (facteur 1, avec une contribution moindre de l'épreuve des paires de mots), celles de CS impliquant les unités syntagmatiques (facteur 2), celles de CS impliquant les unités « mots » (facteur 3) et les variables matricielles non linguistiques (facteur 4). Le modèle ainsi généré permet d'expliquer 74,3% de la variance. Cet état de choses est résumé dans le tableau 8.4.

[74] Les variables dites « latentes » sont des construits théoriques hypothétisés qui ne sont pas mesurés directement, mais qu'on peut inférer à partir de certaines variables observées (Berger, J.-L., 2021).

Tableau 8.4 Contributions des facteurs de l'ACP avec valeur propre supérieure à 1

Contributions des composantes

	Facteurs				
	1	**2**	**3**	**4**	**Unicité**
DerivationFAJOUT	0.847				0.200
DecompositionDSUPP	0.840				0.206
DerivationDAJOUT	0.836				0.172
DecompositionFSUPP	0.823				0.247
ExtractionBaseTotalSUPP	0.675				0.378
PMflexAJOUT	0.596				0.340
PMderivAJOUT	0.579				0.438
PaireMotsSUPP					0.423
CORR_PGRAMM_REP2_S		0.869			0.134
CORR_PGRAMM_REP1_S		0.857			0.107
ID_PGRAMM_REP2_S		0.829			0.142
ID_PGRAMM_REP1_S		0.815			0.137
ID_ERREUR_PAGRAMM_S		0.610			0.342
CORR_ERREUR_PAGRAMM_S		0.554			0.340
CORR_PGRAMM_REP1_M			0.814		0.155
CORR_PGRAMM_REP2_M			0.788		0.185
ID_PGRAMM_REP1_M			0.727		0.228
ID_PGRAMM_REP2_M			0.721		0.199
ID_ERREUR_PAGRAMM_M			0.616		0.375
CORR_ERREUR_PAGRAMM_M			0.607		0.279
MatricesAJOUT				0.842	0.218
ClassificationSUPPR				0.644	0.409

Note. La méthode d'extraction basée sur une valeur propre de 1 a été utilisée, en association avec une rotation 'varimax'. Les contributions inférieures à 0,5 ont été masquées. Le premier facteur représente 23,89% de la variance, le deuxième, 20,97%, le troisième, 19,70%, et le quatrième, 9,74%.

Dans un deuxième temps, nous avons spécifié comme critère d'extraction de composantes une analyse en parallèle, reconnue comme la meilleure[75] (Field *et al.*, 2012). Cette analyse a mis en évidence deux composantes : une regroupant les tests de CM (facteur 2), l'autre les tests de CS (facteur 1), expliquant en tout 64,4 % de la variance, comme en témoigne le tableau 8.5.

[75] L'extraction de composantes est basée sur une analyse en parallèle (*Parallel Analysis* en anglais) dans laquelle des valeurs propres des composantes extraites à partir de données simulées (sans composantes sous-jacentes) sont comparées aux valeurs propres de composantes de la base de données réelles. Les composantes retenues comme composantes principales significatives sont alors celles dont les valeurs propres sont supérieures à celles des données simulées.

Tableau 8.5 Contributions des facteurs de l'ACP avec la méthode d'extraction basée sur une analyse en parallèle

	Facteurs		
	1	2	Unicité
CORR_PGRAMM_REP1_S	0.867		0.208
CORR_PGRAMM_REP2_S	0.845		0.267
ID_PGRAMM_REP1_S	0.842		0.224
ID_PGRAMM_REP2_S	0.830		0.251
ID_PGRAMM_REP2_M	0.700		0.316
ID_PGRAMM_REP1_M	0.688		0.363
ID_ERREUR_PAGRAMM_S	0.661		0.406
CORR_PGRAMM_REP1_M	0.650		0.382
CORR_ERREUR_PAGRAMM_S	0.649		0.380
CORR_PGRAMM_REP2_M	0.645		0.411
CORR_ERREUR_PAGRAMM_M	0.617		0.426
ID_ERREUR_PAGRAMM_M	0.541		0.503
ClassificationSUPPR			0.675
MatricesAJOUT			0.760
DerivationDAJOUT		0.844	0.204
DecompositionDSUPP		0.828	0.272
DerivationFAJOUT		0.809	0.287
DecompositionFSUPP		0.774	0.347
PMflexAJOUT		0.640	0.370
ExtractionBasePrefSUPP		0.623	0.492
PMderivAJOUT		0.606	0.484
ExtractionBaseSuffSUPP		0.574	0.595
PaireMotsSUPP		0.558	0.529

Note. La méthode d'extraction basée sur une analyse en parallèle a été utilisée, en association avec une rotation 'varimax'. Les contributions inférieures à 0,5 ont été masquées. Le premier facteur représente 35,6% de la variance, et le deuxième, 28,7%.

Ces ACP nous permettent d'esquisser un début de réponse à notre première question de recherche, qui est de vérifier si les mesures de CS et de CM peuvent se scinder

en deux constituants – d'ordres linguistique et cognitif – prenant la forme de variables latentes. Il semble que sans apriori sur les données, donc sans le caractère « confirmatoire » des AFC et des SEM, il y a à tout le moins des constituants d'ordre linguistique qui émergent spontanément des données.

Il est à noter que dans ces deux modèles d'ACP, ce sont les unités manipulées dans les épreuves qui semblent ressortir, et qui donnent leur caractère aux différentes composantes. Les opérations ne semblent pas caractériser les facteurs qui émergent dans ces analyses exploratoires. Cette idée se trouve corroborée par l'absence de contribution, ou saturation, pertinente, à un seuil de 0,5 (Berger, J.-L., 2021), des deux variables matricielles – classifications et matrices – aux différentes composantes regroupant les variables de CM et de CS.

8.6 Spécifications des modèles AFC et SEM, et indices d'ajustement

Les analyses factorielles confirmatoires (AFC) appartiennent à la famille des modélisations par équation structurelle (SEM). Ces deux types de modélisations sont utilisées pour tester des hypothèses sur les relations entre des variables. Elles font partie des analyses avec variables latentes, et consistent en l'inclusion de plusieurs indicateurs pour mesurer un phénomène, permettant d'analyser des relations entre variables mesurées et latentes, ou non observées (Bratlie *et al.*, 2022). Le modèle de mesure, correspondant à la partie AFC du modèle, a un intérêt théorique, en sa propension à tester le fonctionnement de dimensions jusque-là hypothétisées,

représentées par des variables latentes (Mehmetoglu et Venturini, 2021). Les variables latentes sont définies par les indicateurs qui les composent. L'ajustement d'un modèle est estimé et permet de déterminer s'il modélise ou représente bien les données.

Pour répondre à notre première question de recherche sur la dimensionnalité de la conscience morphologique et de la conscience syntaxique, nous évaluerons différents modèles grâce à des analyses factorielles confirmatoires (AFC). Ces analyses fourniront ainsi une réponse statistique à notre première question de recherche, qui, nous le rappelons, va comme suit :

- *Les mesures traditionnelles des habiletés en conscience morphologique et en conscience syntaxique peuvent-elles se scinder en deux constituants, l'un linguistique, l'autre cognitif, prenant la forme de variables latentes ?*

Ces analyses nous permettent d'évaluer les modèles de mesure, dont nous faisons l'hypothèse qu'ils rendent compte de la variabilité dans les performances aux construits qui nous intéressent, de les ajuster et de les comparer à d'autres modèles théoriquement valides.

Les modélisations par équation structurelle (SEM) incluent, en plus du modèle de mesure, un modèle structurel, ou modèle de régression structurel, représentant les

relations causales entre les variables latentes du modèle (Kline, 2011). Elles peuvent être considérées comme une combinaison de la régression multiple et de l'analyse factorielle confirmatoire (Kunnan, 1998). Ce type d'analyses permet de vérifier statistiquement des relations hypothétiques entre des construits théoriques représentés par plusieurs variables mesurées (Byrne, 2009; Kline, 2011). Dans une perspective psychométrique, les SEM rendent possible l'exploration de la structure factorielle dans les performances de sujets à des tests ou des questionnaires, mettant en évidence les habiletés évaluées par ces outils de mesure (Byrne, 2009; Ockey et Choi, 2015). L'interprétation du modèle structurel repose sur l'examen des coefficients de régression et des statistiques d'ajustement (Byrne, 2009).

Les SEM permettent ainsi de répondre statistiquement à notre deuxième question de recherche :

- *Dans quelle mesure les variables latentes construites – d'ordre linguistique et cognitif – contribuent-elles aux performances en lecture?*

Les SEM et les AFC requièrent que les modèles soient spécifiés (*input*). Il faut ainsi indiquer, d'une part, quelles variables mesurées construisent (*load on*) les variables latentes hypothétisées – comment les tâches se regroupent en fonction des dimensions exploitées – et d'autre part, quels facteurs sont corrélés (Kline, 2011).

Ces deux types d'analyses nous renseignent (*output*) sur l'ajustement du modèle aux données, en rendant compte des mesures de l'ajustement (*goodness of fit*) des données au modèle théorique (*input*), des coefficients de régression pour les contributions aux variables latentes et des informations diagnostiques (*modification indices*) pour déterminer les défauts dans le modèle (Byrne, 2009; Kline, 2011).

Les étapes de préparation, d'estimation de paramètres, d'ajustement et de comparaison des modèles AFC et SEM sont décrites dans les sections suivantes.

8.6.1 Gestion des données manquantes

Pour que les analyses générées par AFC et SEM soient valides, celles-ci nécessitent des échantillons de données de taille substantielle, en termes d'effectif et de variables mesurées (Byrne, 2009; Kline, 2011). La complexité des collectes de données associées à ce genre d'échantillons occasionne régulièrement des pertes de données; un problème courant qui doit être traité de manière appropriée afin obtenir des estimations de paramètres fiables et des inférences statistiques précises (Byrne, 2009; Loehlin, 2003).

Le caractère foncièrement écologique de notre collecte de données, effectuée en milieu scolaire, occasionne un certain nombre de données manquantes dans la base de données. Ces omissions correspondent le plus souvent à l'absence de certain·es élèves à une séance de collecte de données, pour des raisons logistiques, médicales

ou autres. Certain·es élèves étaient aussi retiré·es momentanément de l'horaire ordinaire de la classe pour obtenir des cours particuliers, ou encore pour troubles du comportement.

Dans ces conditions, les données manquantes peuvent être qualifiées de non aléatoires (*Missing not at random*, MNAR), dans la mesure où ce sous-échantillon d'absent·es est probablement distinct des autres sujets, ou représente un groupe particulier (Bennett, 2001; Saunders *et al.*, 2006). Le choix d'exclure ces données, plutôt que de procéder à d'autres techniques, est risqué. La suppression de cas (*listwise deletion*), qui est certainement répandue, peut entrainer une perte de puissance statistique et des estimations de paramètres biaisées, en particulier si les données ne sont pas manquantes complètement de manière aléatoire (MCAR) (Byrne, 2009). Les données manquantes pour nos analyses SEM, effectuées à l'aide de Jamovi (The jamovi project, 2021) et R (R Core Team, 2020), sont traitées avec une méthode d'estimation à deux étapes robuste, sans imputation, dans laquelle les statistiques de l'échantillon sont estimées (vecteur moyen, matrice de variance-covariance) à l'aide d'un algorithme EM (*Expectation Maximization*) à partir de l'échantillon et utilisées dans l'analyse subséquente, comme si les données étaient complètes (Rosseel, 2012). Cette méthode est préférée lorsque les données manquantes sont MNAR (Fernandez-Garcia *et al.*, 2018).

8.6.2 Méthode d'estimation

Les méthodes d'estimation de paramètres de type Maximum de vraisemblance (ML) permettent de déterminer les valeurs des paramètres qui maximisent la vraisemblance de la distribution des données observées conditionnellement aux variables latentes dans les modèles d'équations structurelles (SEM) (Kline, 2011; Kunnan, 1998). Le caractère robuste de la méthode d'estimation permet d'estimer les paramètres du modèle en présence de données qui ne respectent pas les hypothèses de normalité et d'homoscédasticité (Loehlin, 2003). L'utilisation de la méthode Maximum de Vraisemblance Robuste (*Robust maximum Likelihood*, MLMVS) est donc motivée par l'observation de légères asymétries négatives pour certaines variables, plus marquées lorsqu'elles sont divisées en fonction du niveau scolaire. Cette procédure est notamment recommandée par Kunnan (1998) dans son ouvrage sur les recherches sur les évaluations linguistiques.

8.6.3 Indices d'ajustement et informations sur les modèles

Suivant les recommandations de Byrne (2009) et de Xuelian *et al.* (2019) et les procédures suivies par Zhang, L. (2018), nous avons retenu les critères suivants pour évaluer la qualité d'ajustement des modèles AFC et SEM aux données (*model fit indices*) : RMSEA, khi carré normé (X^2/df), CFI, TLI, SRMR, AIC et GFI.

Les normes permettant de statuer sur la fiabilité et la crédibilité des modèles évalués sont régulièrement remises en question dans la littérature scientifique dédiée à ces

méthodes d'analyses statistiques (Byrne, 2009). Nous retiendrons les critères les plus couramment utilisés. Les CFI et TLI doivent être au-dessus de 0,90 pour un ajustement assez bon, ou au-dessus de 0,95 pour un ajustement bon (Xuelian *et al.*, 2019). Les RMSEA doivent être sous la barre des 0,06, et leurs intervalles de confiance doivent être minces. Les valeurs entre 0,06 et 0,08 représentent des erreurs d'approximation raisonnables, celles comprises entre 0,08 et 0,10 indiquent un ajustement pauvre, et celles supérieures à 0,10, un ajustement aux données médiocre (Byrne, 2009). Il est à noter que lorsque la taille de l'échantillon est faible, la RMSEA et le TLI ont tendance à mal évaluer et à rejeter les modèles de la population réelle. Les valeurs du SRMR doivent être inférieures à 0,10 (Byrne, 2009). Le khi deux normé (X^2/df) doit être de moins de 3 (Zhang, L., 2018). Les valeurs de GFI les plus près de 1 indiquant quant à elles un bon ajustement du modèle (Byrne, 2009). Le GFI est fortement affecté par la taille de l'échantillon : plus l'effectif est réduit, plus il tend à baisser (Xuelian *et al.*, 2019).

Les SEM donnent également de précieuses informations sur le caractère prédictif du modèle de mesure à travers le R^2 et les coefficients de régression standardisés (bêtas) entre variables latentes endogènes et exogènes.

8.6.4 Multicolinéarité

Le tableau 8.4, présentant les coefficients de corrélation entre les différentes variables, fait état d'une corrélation très forte (r < 0,90) entre les variables corr_pagramm_rep1_s et corr_pagramm_rep2_s, indicative d'une situation de multicolinéarité, qui peut entrainer des estimations instables, une surestimation des coefficients de régression, ainsi qu'une augmentation des erreurs standards, rendant difficile l'interprétation des résultats (Xuelian *et al.*, 2019). Nous avons donc supprimé l'une des deux variables, soit la variable corr_pagramm_rep2_s., suivant une stratégie recommandée par Kline (2011), afin de mener à bien les analyses subséquentes.

8.6.5 Comparaison de modèles

La comparaison de modèles hiérarchiques, ou imbriqués – un modèle imbriqué dans un autre peut par exemple comporter un paramètre libre de moins – est évaluée à l'aide du test de différence de khi carré ($\Delta\chi^2$)[76] (Byrne, 2009; Kline, 2011). Une mesure plus simple de comparaison de modèles est également proposée dans la littérature sur les SEM : la différence de CFI entre modèles (ΔCFI) (Byrne, 2009). Celle-ci est réputée significative si elle est supérieure au seuil de > 0,01 (Byrne, 2009; Cheung et Rensvold, 2002).

[76] La formule pour le test de différence de khi carré est la suivante :
$\Delta\chi^2 = \chi^2 1 - \chi^2 2$
où $\Delta\chi^2$ est la différence des valeurs de khi carré entre les deux modèles, $\chi^2 1$ est la valeur de khi carré pour le modèle avec le moins de paramètres libres, et $\chi^2 2$ est la valeur de khi carré pour le modèle avec le plus de paramètres libres.

Pour évaluer la différence de qualité entre modèles non hiérarchiques, il est possible de comparer leurs valeurs associées au critère d'information d'Akaike (désormais AIC) (Kline, 2011), qui constitue un autre indice de comparaison de modèles[77]. Les petites valeurs d'AIC sont indicatives de bons ajustements des modèles aux données, et les grandes valeurs, de mauvais ajustements (Loehlin, 2003).

8.7 Plan d'analyse des modèles SEM et AFC appliquées aux variables de CM et de CS

Dans le cadre de nos analyses AFC et SEM, nous avons conçu des modèles en adéquation avec les résultats de nos ACP et la littérature exposée dans notre cadre théorique de la partie 1. Nous avons également adopté une stratégie d'évaluation des modèles en question basée sur le repérage de changements dans les indices d'ajustement et dans les coefficients standardisés entre les différents modèles, en conformité avec les pratiques les plus couramment utilisées dans le domaine et décrites dans la section 8.6.3 (Byrne, 2009; Kline, 2011; Levesque *et al.*, 2017; Ockey et Choi, 2015; Xuelian *et al.*, 2019; Zhang, L., 2018).

Les épreuves de conscience morphologique et de conscience syntaxique sont tour à tour soumises à deux types d'analyses : d'abord, les analyses factorielles

[77] La formule la plus simple permettant de calculer l'AIC est la suivante :
$AIC = \chi^2+2q$
où q est le nombre de paramètres libres du modèle.

confirmatoires (AFC), ensuite, les analyses par SEM, dans lesquelles figurent les modèles structurels. Les modèles testés comprennent des modèles unidimensionnels ainsi que des modèles bidimensionnels, les premiers impliquant des opérations et les autres impliquant des unités. Ainsi, pour chacun des construits, une série de quatre AFC sont effectuées : un modèle unidimensionnel, un modèle bidimensionnel mettant en œuvre les opérations, un autre mettant en œuvre les opérations avec en plus les variables matricielles, et un modèle bidimensionnel mettant en œuvre les unités.

Au regard des résultats obtenus pour les AFC, nous choisissons de ne sélectionner que les modèles les plus pertinents pour les analyses par SEM (Kunnan, 1998). Les analyses SEM appliquées aux données de CM et de CS constituent le prolongement des modèles AFC testés pour les deux construits. C'est donc par un souci de parallélisme avec les AFC effectuées, et pour des raisons de pertinence théorique que nous présentons les indices d'ajustement de tous ces modèles (unidimensionnels et bidimensionnels). Par la suite, seuls les modèles travaillés par SEM ayant fourni un ajustement adéquat aux données sont sélectionnés et font l'objet d'analyses commentées, notamment au regard des contributions[78] des variables mesurées aux

[78] Dans une analyse factorielle confirmatoire, les coefficients de régression, ou bêtas, permettent d'interpréter les charges factorielles des variables mesurées, c'est-à-dire comment les variables observées sont associées aux facteurs latents sous-jacents. Ils indiquent la force, la direction et l'importance de ces relations. Dans une analyse SEM complète, c'est-à-dire avec modèle structurel, les coefficients de régression sont également indiqués entre les variables latentes exogènes et endogènes. Par exemple, avec une variable latente exogène (ici, variable indépendante) dont la valeur bêta pour représenter la relation avec une variable endogène (ici, variable dépendante) est estimée à 0,50, cela signifierait que chaque augmentation d'une unité d'écart-type de la variable exogène entraîne une augmentation de 0,50 unité d'écart type dans la variable exogène.

variables latentes, et des variables latentes exogènes vers la variable latente endogène (la compréhension).

8.8 Analyses appliquées aux variables de CM

Nous présentons en 8.8.1 les AFC pour les épreuves de conscience morphologique, et en 8.8.2, les SEM appliqués à ces mêmes épreuves.

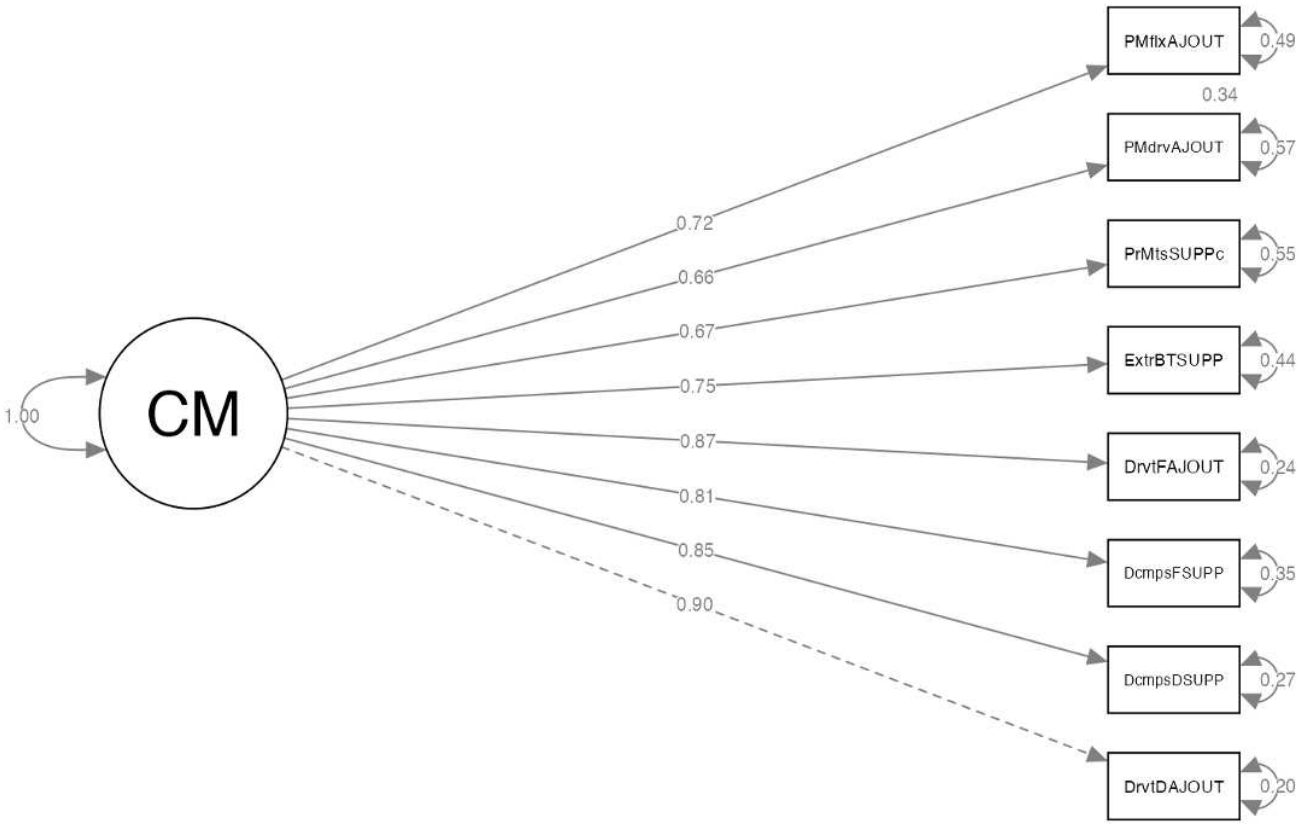
8.8.1 Analyses factorielles confirmatoires appliquées à la CM

Il est à noter que le modèle unidimensionnel de la CM et le modèle bidimensionnel impliquant les opérations dans les épreuves de CM ont tous deux bénéficié de l'ajout d'une covariance entre les termes d'erreurs des variables PMderivAJOUT~~PMflexAJOUT, un indice de modification qui était systématiquement recommandé dans le *output* du logiciel, en plus d'être théoriquement valide. Cette covariance n'a pas été ajoutée au modèle bidimensionnel impliquant les unités, étant donné qu'il était opposé à l'hypothèse théorique sous-jacente dans ce modèle.

8.8.1.1 Modèle unidimensionnel de la CM

Nous avons d'abord spécifié un modèle unidimensionnel de la conscience morphologique (figure 8.1), dans lequel toutes les tâches se rapportent à une seule dimension, à savoir la conscience morphologique. Le modèle s'adapte bien aux

données (X^2 : 32, df : 19, RMSEA : 0,06, CFI : 0,987, TLI : 0,982, SRMR : 0,028, GFI : 0,952, AIC : 82), notamment avec l'ajout d'une covariance entre les erreurs (PMderivAJOUT~~PMflexAJOUT). Tous les liens entre variables mesurées et variables latentes sont significatifs ($p<0,001$).



Note. PMflxAJOUT : Choix de pseudomots suffixés, morphèmes flexionnels; PMdrvAJOUT : Choix de pseudomots suffixés, morphèmes dérivationnels; PrMtsSUPPc : Paires de mots, morphèmes dérivationnels; ExtrBTSUPP : Extraction de la base, morphèmes dérivationnels; DrvtFAJOUT : Dérivation, morphèmes flexionnels; DcmpsFSUPP : Décomposition, morphèmes flexionnels; DcmpsDSUPP : Décomposition, morphèmes dérivationnels; DrvtDAJOUT : Dérivation, morphèmes dérivationnels

Figure 8.1 AFC du modèle unidimensionnel de la CM.

Les contributions des variables DérivationDAJOUT et DérivationFAJOUT sont les plus élevées (respectivement β = 0,90 et β = 0,87), suivies des variables DécompositionDSUPP et DécompositionFSUPP (respectivement β = 0,85 et β = 0,81). Les autres variables mesurées contribuent également de manière significative à la construction de la variable latente représentant la CM. Leurs contributions (β) s'échelonnent de 0,67 à 0,75.

8.8.1.2 Modèle bidimensionnel de la CM impliquant les unités

Nous avons spécifié un premier modèle bidimensionnel impliquant les unités linguistiques à l'œuvre dans les épreuves de CM, dans lequel nous supposons que les types de morphèmes (flexionnels et dérivationnels) mobilisés dans les tâches représentent les différences individuelles dans la qualité des réponses. Ce modèle présente des indices d'ajustement acceptables (X^2 : 52,9, df : 19, RMSEA : 0,097, CFI : 0,967, TLI : 0,952, SRMR : 0,035, GFI : 0,929, AIC : 102,9). Cependant, celui-ci est significativement moins bien ajusté aux données que le modèle unidimensionnel, comme en témoigne le tableau 8.7. Une analyse plus fine des données permet de rendre compte d'un problème plus grave : la présence de corrélations supérieures à 1 entre les variables latentes. Cet état de choses indique probablement une spécification inadéquate du modèle aux données (Byrne, 2009; Kunnan, 1998), et peut être le fruit d'une trop grande multicolinéarité entre les variables construites.

8.8.1.3 Modèle bidimensionnel de la CM impliquant les opérations

Conformément à nos objectifs de recherche, nous avons spécifié un second modèle bidimensionnel, mettant cette fois-ci en œuvre les types d'opérations (X^2 : 29,5, df : 18, RMSEA : 0,058, CFI : 0,989, TLI : 0,983, SRMR : 0,025, GFI : 0,96, AIC : 81,5). Dans ce cas-ci également, la présence de corrélations au-dessus de 1 pour les deux variables latentes rend compte d'un problème dans les spécifications du modèle (Byrne, 2009). L'ajout des épreuves non linguistiques dans les modèles avec opérations n'a pas permis d'améliorer l'ajustement (X^2 : 78,9, df : 33, RMSEA : 0,091, CFI : 0,955, TLI : 0,939, SRMR : 0,051, GFI : 0,887, AIC : 142,9), certains indices révélant même un moins bon ajustement aux données.

Les indices d'ajustement de tous les modèles testés figurent dans le tableau 8.6, et le tableau 8.7 fait état des valeurs relatives aux différences d'ajustement et des tests de khi carré.

Tableau 8.6 Indices d'ajustement pour les modèles des AFC appliquées aux épreuves de CM

Modèle	NPAR	X^2	df	*p* X^2	X^2/df	RMSEA	RMSEA CI 95%	*p* RMSEA	CFI	TLI	SRMR	GFI	AIC
CM Opérations	26	29,5	18	0,043	1,639	0,058	0,011-0,095	0,328	0,989	0,983	0,025	0,96	81,5
CM Opérations et matrices	32	78,9	33	<0,001	2,391	0,091	0,065-0,117	0,006	0,955	0,939	0,051	0,887	142,9
CM Unités	25	52,9	19	<0,001	2,784	0,097	0,067-0,129	0,007	0,967	0,952	0,035	0,929	102,9
CM Unidimensionnel	25	32	19	<0,032	1,684	0,06	0,018-0,095	0,295	0,987	0,982	0,028	0,952	82

Note. Les indices en vert indiquent un excellent ajustement aux données, les indices en jaune représentent un ajustement acceptable, et les indices en rouge marquent un mauvais ajustement.

Tableau 8.7 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles des AFC appliquées aux épreuves de CM

AFC	ΔX^2	Δdf	*p*	ΔRMSEA	ΔCFI	ΔTLI	ΔSRMR	ΔGFI	ΔAIC
CM Opérations - CM Unidimensionnel	-2,5	-1	< .00001	-0,002	0,002	0,001	-0,003	0,008	-0,5
CM Opérations et matrices - CM Opérations	49,4	15	< .00001	0,033	-0,034	-0,044	0,026	-0,073	61,4
CM Unités - CM Unidimensionnel	20,9	0	< .00001	0,037	-0,020	-0,030	0,007	-0,023	20,900

Comme mentionné auparavant, les indices d'ajustement nous renseignent sur la qualité d'ajustement des modèles statistiques hypothétisés. Nous constatons à leur inspection que les modèles relatifs à la conscience morphologique s'ajustent tous plutôt bien aux données, comme l'indiquent la plupart des indices, avec notamment les CFI de plus de 0,95, les TLI de plus de 0,90 et les SRMR de moins de 0,10, à l'exception notable du modèle Opération et matrices. Les valeurs du RMSEA de ces modèles impliquant les opérations avec variables matricielles sont hors de la plage recommandée d'acceptabilité (< 0,06 à 0,10), pointant vers une moins bonne spécification de ces modèles. En effet, le modèle Opérations avec variables matricielles présente des indices d'ajustement moins bons que le modèle CM opérations, comme indiqué dans le tableau 8.7.

Vu les différences d'ajustement et de khi carré significatives entre les modèles unidimensionnels et les modèles mettant en œuvre deux dimensions, et compte tenu surtout des corrélations au-dessus de 1 remarquées entre les variables latentes de tous les modèles bidimensionnels évalués, nous devons conclure que le modèle unidimensionnel est le mieux à même de représenter les mesures de CM.

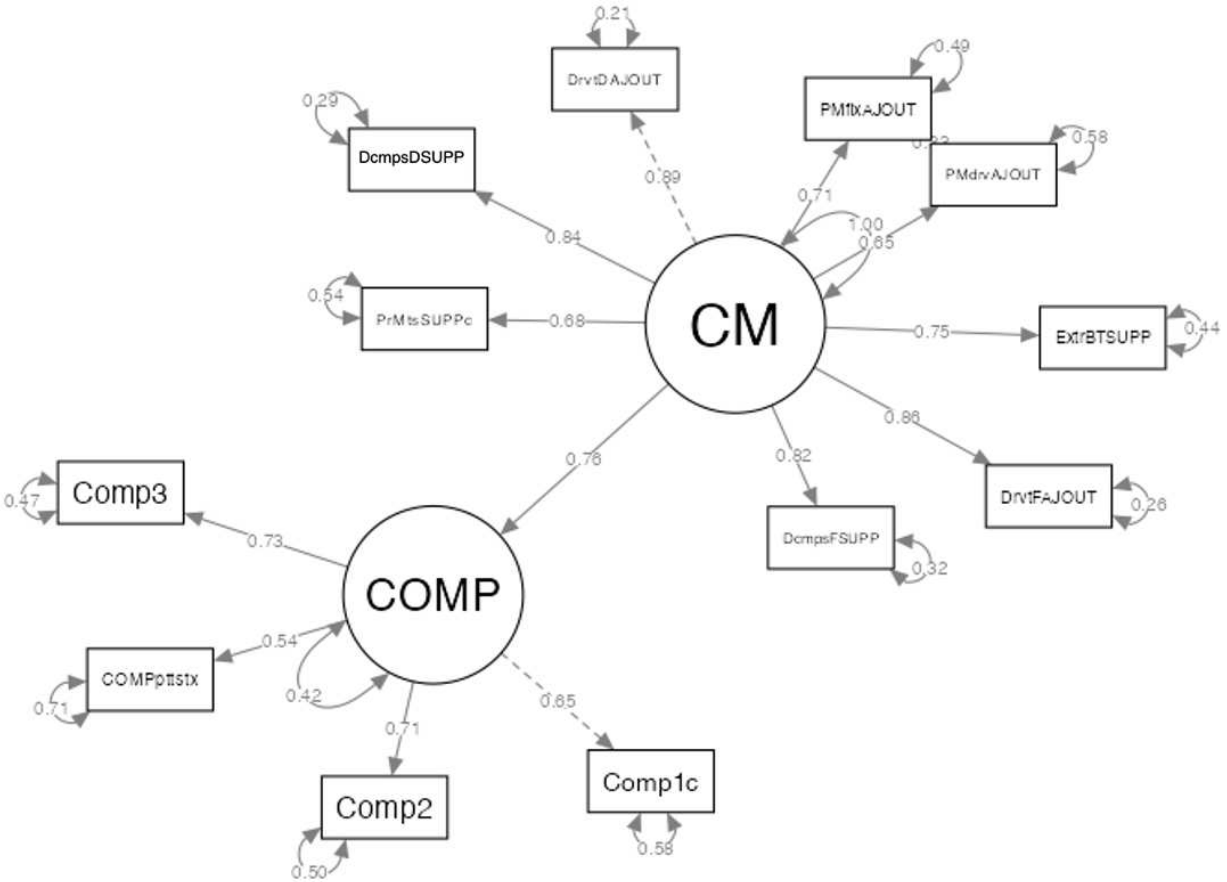
8.8.2 Modélisations par équation structurelle (SEM) appliquées à la CM

Le modèle unidimensionnel, avec opérations et avec unités fait, dans cette section, l'objet d'analyses SEM comparées. Partiellement redondantes au regard des AFC présentées précédemment, elles présentent l'avantage de matérialiser

statistiquement notre deuxième question de recherche. Nous les abordons donc dans la suite de cette section.

8.8.2.1 Modèle unidimensionnel de la CM

Le modèle CM unidimensionnel, avec l'inclusion d'une covariance (PMderivAJOUT ~~ PMflexAJOUT), est ajusté. La variable latente représentant la version unitaire de la CM contribue significativement à expliquer la compréhension ($\beta = 0,76$).



. Note. COMPpttstx : Compréhension textes courts; Comp1c : Compréhension littérale, texte long; Comp2 : Inférences, texte long; Comp3 : Liens entre les deux premiers niveaux, texte long; PMflxAJOUT : Choix de pseudomots suffixés, morphèmes flexionnels; PMdrvAJOUT : Choix de pseudomots suffixés, morphèmes dérivationnels; PrMtsSUPPc : Paires de mots, morphèmes dérivationnels; ExtrBTSUPP : Extraction de la base, morphèmes dérivationnels; DrvtFAJOUT : Dérivation, morphèmes flexionnels; DcmpsFSUPP : Décomposition, morphèmes flexionnels; DcmpsDSUPP : Décomposition, morphèmes dérivationnels; DrvtDAJOUT : Dérivation, morphèmes dérivationnels.

Figure 8.2 Représentation graphique du modèle SEM avec les épreuves de conscience morphologique comme construit unidimensionnel

En ce qui a trait à la variable latente représentant la compréhension en lecture, les contributions des variables Comp2 (β = 0,71) et Comp3 (β = 0,73) sont supérieures à celles des variables COMPpetits textes (β = 0,54) et Comp1 (β = 0,65). Le modèle de mesure explique 57,7% de la compréhension telle qu'elle est mesurée dans la présente recherche.

8.8.2.2 Modèle bidimensionnel de la CM impliquant les opérations

Le modèle SEM pour la CM impliquant les opérations a bénéficié de l'ajout d'une covariance (PMderivAJOUT~~PMflexAJOUT), recommandée par le logiciel dans les indices de modification et valide sur le plan théorique.

À l'instar de ce qui avait été observé dans les AFC, les modèles font état de corrélations au-dessus de 1 entre les deux variables latentes, ce qui indique une mauvaise spécification du modèle (Byrne, 2009), comme mentionné auparavant. Il n'est donc pas pertinent de commenter les indices d'ajustement et les contributions du modèle plus avant.

8.8.2.3 Modèles bidimensionnels de la CM impliquant les unités

Les indices d'ajustement pour les modèles impliquant les morphèmes dérivationnels et flexionnels indiquent que ceux-ci ne sont que partiellement ajustés aux données. Aucune des deux variables latentes ne prédit significativement la variable

dépendante, et les indices de modification proposés ne sont pas valides sur le plan théorique. À nouveau, et en conformité avec l'AFC correspondante, le modèle implique des corrélations au-dessus de 1 entre les variables latentes représentant les épreuves avec morphèmes flexionnels et celles qui comprennent des morphèmes dérivationnels, ce qui rend caduque une description plus fine du modèle. Notre interprétation de ce résultat est que les variables latentes présentent une multicolinéarité, sur le plan statistique, c'est-à-dire qu'elles se recoupent, ou mesurent des variables trop similaires pour bien décrire les données.

Le tableau 8.8 résume l'ensemble de ces informations et fournit les indices d'ajustement pour tous les modèles testés. Le tableau 8.9 fait état de différences d'ajustement et des valeurs des tests de khi carré entre les modèles, les différences des indices AIC et les éléments concernant la multicoliénarité des variables latentes des deux modèles bidimensionnels confirmant que le modèle unidimensionnel est à privilégier.

Tableau 8.8 Indices d'ajustement pour les modèles des SEM appliquées aux épreuves de CM

Modèle	NPAR	X^2	df	*p* X^2	X^2/df	RMSEA	RMSEA CI 95%	*p* RMSEA	CFI	TLI	SRMR	R^2	GFI	AIC
CM Opérations	40	109,3	50	<0,001	2,186	0,083	0,061-0,104	0,007	0,952	0,936	0,057	0,576	0,845	189,3
CM Unités	39	129,1	51	<0,001	2,531	0,094	0,074-0,114	<0,001	0,936	0,917	0,059	0,52	0,831	207,1
CM Unidimensionnel	38	112,2	52	<0,001	2,158	0,082	0,061-0,102	0,008	0,951	0,937	0,059	0,577	0,84	188,2

Note. Les indices en vert indiquent un excellent ajustement aux données, les indices en jaune représentent un ajustement acceptable, et les indices en rouge marquent un mauvais ajustement.

Tableau 8.9 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles des SEM appliquées aux épreuves de CM

SEM	ΔX^2	Δdf	*p*	ΔRMSEA	ΔCFI	ΔTLI	ΔSRMR	ΔGFI	ΔAIC
CM Opérations - CM Unidimensionnel	-2,9	-2	< .00001	0,001	0,001	-0,001	-0,002	0,005	1,1
CM Unités - CM Unidimensionnel	16,9	-1	< .00001	0,012	-0,015	-0,02	0	-0,009	18,9

8.8.3 Synthèse des résultats sur la conscience morphologique

Les résultats des AFC indiquent que le modèle unidimensionnel représente le mieux les données, un état de choses qui soutient une conception unitaire, unifiée, de la conscience morphologique, ainsi que la cohérence interne des instruments servant à la mesurer.

Quant aux SEM, ils nous renvoient également l'image d'une conscience morphologique unidimensionnelle. Les résultats nous renseignent aussi sur le caractère prédictif du construit. En effet, la variable latente représentant la CM prédit significativement la variable Compréhension.

8.9 Analyses appliquées aux variables de CS

Les analyses appliquées aux épreuves de conscience syntaxique sont tout à fait analogues à celles de conscience morphologique, et cherchent à en vérifier la dimensionnalité (8.9.1), puis à vérifier quel modèle permet de prédire la compréhension (8.9.2).

8.9.1 Analyses factorielles confirmatoires appliquées à la CS

Les modèles des AFC qui se rapportent à la CS sont le miroir de ceux qui ont été présentés en 8.8.1 et qui se rapportaient à la CM, soit des modèles unidimensionnels,

et des modèles bidimensionnels divisés en fonction des variables hypothétisées : unités et opérations. L'ajout de covariances pour améliorer les modèles est mentionné dans les sections dédiées aux modèles.

8.9.1.1 Modèle unidimensionnel de la CS

Le modèle unidimensionnel de la CS ne convainc pas, et ce, malgré l'ajout de deux covariances pour les termes d'erreurs (ID_ERREUR_PAGRAMM_Mc~~CORR_ERREUR_PAGRAMM_M, ID_PGRAMM_REP1~~ID_PGRAMM_REP2_S). Les mesures de la CS sont donc mal représentées par un modèle unidimensionnel, dans lequel la CS constitue un construit unitaire, contrairement à ce que nous avions constaté au regard des épreuves de CM.

8.9.1.2 Modèle bidimensionnel de la CS impliquant les opérations

Le modèle impliquant des opérations d'addition et de suppression dans les épreuves de CS n'est pas bien ajusté aux données, malgré l'ajout d'une covariance entre ID_PGRAMM_REP1_S ~~ ID_PGRAMM_REP2_S, recommandée par les *output* des indices de modification. Les valeurs négatives des GFI que présentent les deux modèles bidimensionnels avec opérations – avec et sans matrices – sont étonnantes, bien que recensées dans la littérature (Cheung et Rensvold, 2002)[79]. Les autres

[79] Il semble que la présence de GFI négatifs soit favorisée dans des conditions où la taille de l'échantillon est de 200 sujets et où des méthodes d'estimation de maximum de vraisemblance sont utilisées (Byrd, 2008), ce qui est le cas pour notre étude.

indices d'ajustement étant bien en deçà des normes, nous tirons la conclusion que les modèles impliquant les opérations ne sont malheureusement pas convaincants, et ne permettent pas de décrire adéquatement la variation au sein de l'échantillon.

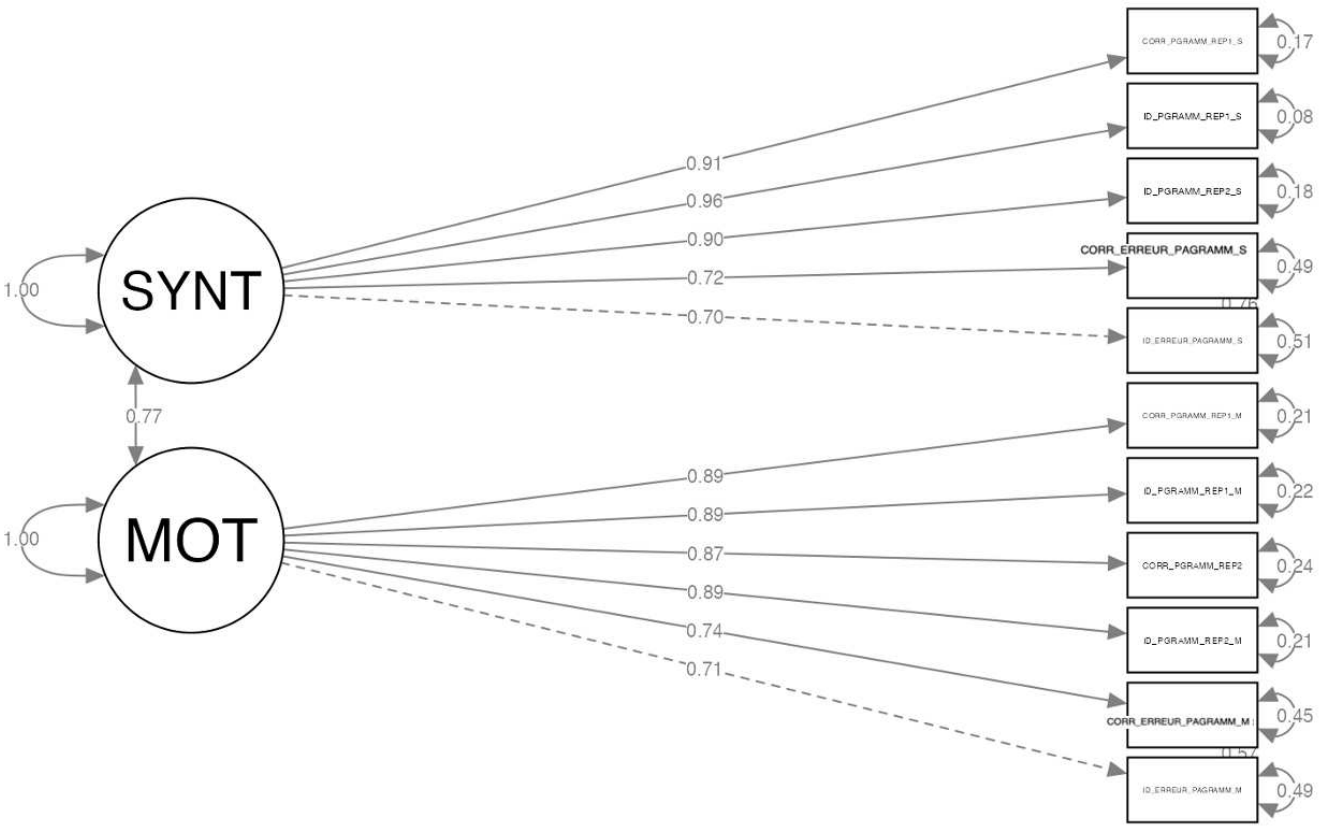
À l'instar des modèles impliquant des opérations dans les AFC dédiés à la CM, les modèles de CS révèlent des corrélations au-dessus de 1 pour variables latentes, tant dans les modèles sans variables matricielles non linguistiques que les modèles avec ces variables. Ainsi, les modèles mettant en œuvre les opérations, avec ou sans matrices, font état d'écueils similaires à ceux retrouvés dans les AFC dédiées aux épreuves de CM au regard des corrélations au-dessus de 1 pour les variables latentes, en plus de présenter des indices d'ajustement insatisfaisants.

8.9.1.3 Modèle bidimensionnel de la CS impliquant les unités

Le modèle AFC spécifié impliquant deux unités syntaxiques, ou grammaticales, dans le cas de notre étude, s'adapte assez bien aux données (X^2 : 129,6, df : 41, RMSEA : 0,109, CFI : 0,958, TLI : 0,943, SRMR : 0,062, GFI : 0,784, AIC : 201,6), notamment avec l'ajout d'une covariance entre les erreurs (PMderivAJOUT~~PMflexAJOUT). Tous les liens entre variables mesurées et variables latentes sont significatifs ($p<0,001$). Les AFC effectuées ont permis d'extraire des facteurs d'ordre linguistique dans les données, et confirment la pertinence des variables latentes hypothétisant la présence de construits sous-jacents représentant deux unités syntaxiques de tailles différentes pour les épreuves de CS – l'une se rapportant à des mots qui représentent

des classes grammaticales, et l'autre se rapportant à des groupes ayant des fonctions syntaxiques. La figure 8.3 permet de visualiser les liens entre les variables observées et les variables latentes.

Les contributions des variables d'identification de l'erreur dans des phrases agrammaticales, tant dans les mots que dans les syntagmes, sont les plus faibles dans le modèle (ID_ERREUR_PAGRAMM_M : β = 0,71; ID_ERREUR_PAGRAMM_S : β =0,70). Les contributions des variables mesurées se référant aux items de réplication de l'erreur (sur une première phrase grammaticale et sur une deuxième) sont les plus élevées (s'échelonnant de β = 0,90 à β = 0,960 pour la variable latente SYNT, et de β = 0,87 à β = 0,89 pour la variable latente MOT). Toutes les variables mesurées contribuent de manière significative à la construction des deux variables latentes représentant la CS ($p < 0,001$).



Note. ID_ERREUR_PAGRAMM_M : Identification de l'erreur dans une phrase agrammaticale (mot); CORR_ERREUR_PAGRAMM_M : Correction de l'erreur dans phrase agrammaticale (mot); ID_PGRAMM_REP1_M; CORR_PGRAMM_REP1_M : Identification dans une phrase grammaticale, réplication 1 (mot); ID_PGRAMM_REP2_M : Identification dans une phrase grammaticale, réplication 2 (mot); CORR_PGRAMM_REP2_M : Correction dans une phrase grammaticale, réplication 2 (mot); ID_ERREUR_PAGRAMM_S : Identification de l'erreur dans phrase agrammaticale (syntagme); CORR_ERREUR_PAGRAMM_S : Correction de l'erreur dans phrase agrammaticale (syntagme); ID_PGRAMM_REP1_S : Identification dans une phrase grammaticale, réplication 1 (syntagme), CORR_PGRAMM_REP1_S : Correction dans une phrase grammaticale, réplication 1 (syntagme); ID_PGRAMM_REP2_S : Identification dans une phrase grammaticale, réplication 2 (syntagme) SYNT : syntagme.

Figure 8.3 AFC du modèle bidimensionnel de la CS impliquant les unités.

Les indices d'ajustement des différents modèles figurent dans le tableau 8.10. L'étude du tableau permet de confirmer que le modèle CS Unités est le seul à bien représenter les données; un constat qu'il est également possible d'inférer à partir du tableau 8.11, qui présente les différences d'ajustement et les valeurs des tests khi carré. En effet, le test de khi carré permet de statuer sur le meilleur ajustement du modèle CS Unités par rapport au modèle unidimensionnel (ΔX^2 : -252,4, Δdf : -1, $p < 0{,}0001$).

Tableau 8.10 Indices d'ajustement pour les modèles AFC appliquées aux épreuves de CS

Modèle	NPAR	X^2	df	*p* X^2	X^2/df	RMSEA	RMSEA CI 95%	*p* RMSEA	CFI	TLI	SRMR	GFI	AIC
CS Opérations	35	463	42	<0,001	11,024	0,235	0,216-0,254	<0,001	0,798	0,736	0,071	-0,554	533
CS Opérations et matrices	41	460	63	<0,001	7,302	0,196	0,179-0,213	<0,001	0,799	0,751	0,196	-0,316	542
CS Unités	36	129,6	41	<0,001	3,161	0,109	0,088-0,130	<0,001	0,958	0,943	0,062	0,784	201,6
CS Unidimensionnel	35	382	42	<0,001	9,095	0,211	0,192-0,230	<0,001	0,837	0,787	0,063	0,195	452

Note. Les indices en vert indiquent un excellent ajustement aux données, les indices en jaune représentent un ajustement acceptable, et les indices en rouge marquent un mauvais ajustement.

Tableau 8.11 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles AFC appliquées aux épreuves de CS

AFC	ΔX^2	Δdf	*p*	ΔRMSEA	ΔCFI	ΔTLI	ΔSRMR	ΔGFI	ΔAIC
CS Opérations - CS Unidimensionnel	81	0	< .001	0,024	-0,039	-0,051	0,008	-0,749	81,00
CS Opérations et matrices - CS Opérations	-3	21	> .05	-0,039	0,001	0,015	0,125	0,238	9,00
CS Unités - CS unidimensionnel	-252,4	-1	< .001	-0,102	0,121	0,156	-0,001	-0,589	-250

8.9.2 Modélisations par équations structurelles (SEM) appliquées à la CS

Les résultats des AFC précédemment exposées justifient d'aller de l'avant avec les analyses SEM impliquant les dimensions linguistiques à l'œuvre dans les performances en compréhension.

8.9.2.1 Modèle bidimensionnel CS impliquant les unités

Le modèle, qui inclut deux covariances (ID_ERREUR_PAGRAMM_S ~~ CORR_ERREUR_PAGRAMM_Sc, ID_ERREUR_PAGRAMM_Mc ~~ CORR_ERREUR_PAGRAMM_M, s'ajuste bien aux données (X^2 : 184, df : 85, RMSEA : 0,083, CFI : 0,953, TLI : 0,942, SRMR : 0,066, GFI : 0,749, AIC : 284). Les contributions des deux variables latentes relatives aux unités syntaxiques – mots représentant différentes classes grammaticales et syntagmes, ou groupes représentant des fonctions – ont un impact significatif sur la compréhension en lecture. Toutes deux présentent des coefficients de régression standardisés positifs et similaires (respectivement 0,38 pour la variable syntagme et 0,41 pour la variable mot), comme la figure 8.4 l'indique. Les deux variables latentes présentent une covariance de 0,74, indiquant qu'elles évoluent ensemble. Le modèle spécifié avec deux variables exogènes prédit 53,7 % de la variance totale de la variable endogène (R^2=0,537).

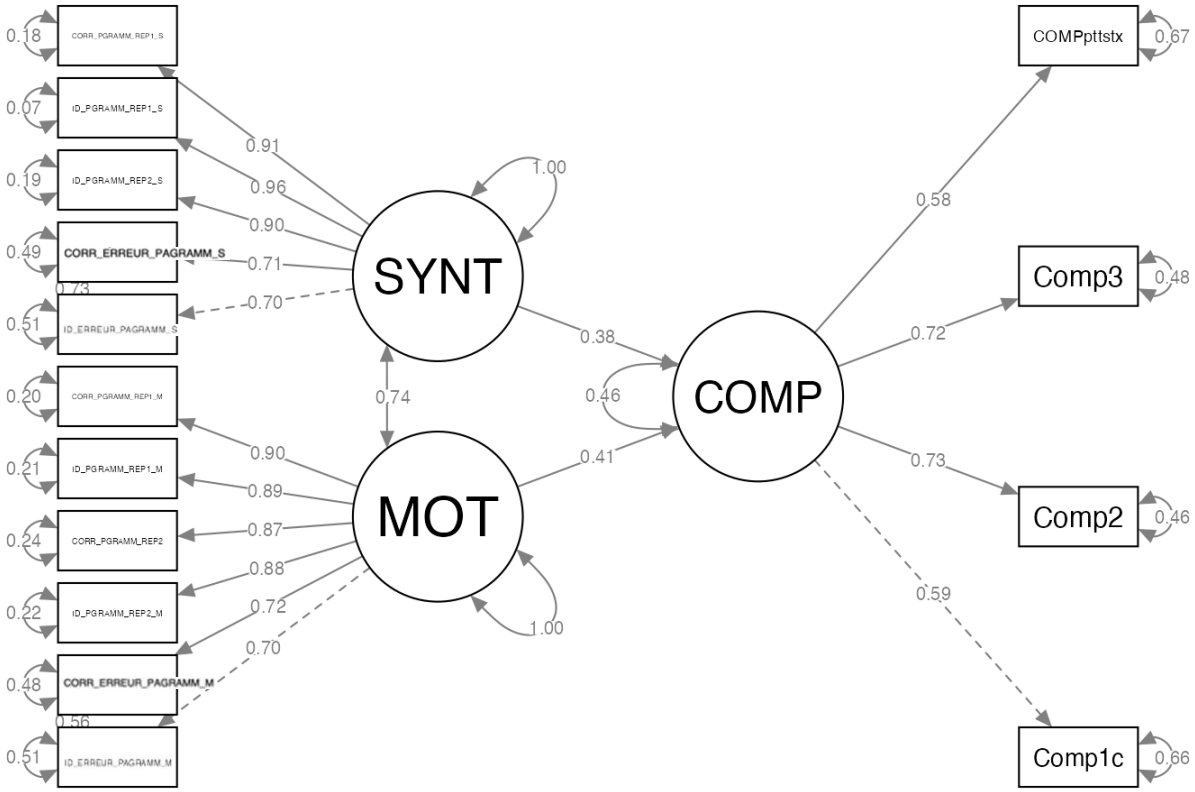


Figure 8.4 Représentation graphique du modèle SEM avec les épreuves de conscience syntaxique comme construit bidimensionnel, divisé en fonction des unités mises en œuvre

Note. Les modèles impliquent des relations de covariances entre les variables suivantes : ID_ERREUR_PAGRAMM_S~~CORR_ERREUR_PAGRAMM_Sc et ID_ERREUR_PAGRAMM_Mc~~CORR_ERREUR_PAGRAMM_M. COMPpttstx : Compréhension textes courts; Comp1c : Compréhension littérale, texte long; Comp2 : Inférences, texte long; Comp3 : Liens entre les deux premiers niveaux, texte long;ID_ERREUR_PAGRAMM_M : Identification de l'erreur dans une phrase agrammaticale (mot); CORR_ERREUR_PAGRAMM_M : Correction de l'erreur dans phrase agrammaticale (mot); ID_PGRAMM_REP1_M; CORR_PGRAMM_REP1_M : Identification dans une phrase grammaticale, réplication 1 (mot); ID_PGRAMM_REP2_M : Identification dans une phrase grammaticale, réplication 2 (mot); CORR_PGRAMM_REP2_M : Correction dans une phrase grammaticale, réplication 2 (mot); ID_ERREUR_PAGRAMM_S : Identification de l'erreur dans une phrase agrammaticale (syntagme); CORR_ERREUR_PAGRAMM_S : Correction de l'erreur dans une phrase agrammaticale (syntagme); ID_PGRAMM_REP1_S : Identification dans une phrase grammaticale, réplication 1 (syntagme), CORR_PGRAMM_REP1_S : Correction dans une phrase grammaticale, réplication 1 (syntagme); ID_PGRAMM_REP2_S : Identification dans une phrase grammaticale, réplication 2 (syntagme) SYNT : syntagme.

Tableau 8.12 Indices d'ajustement pour le modèle de SEM appliquée aux épreuves de CS

Modèle	NPAR	X^2	df	*p* X^2	X^2/df	RMSEA	RMSEA CI 95%	*p* RMSEA	CFI	TLI	SRMR	R^2	GFI	AIC
CS Unités	50	184	85	<0,001	2,165	0,083	0,067-0,100	<0,001	0,953	0,942	0,066	0,537	0,749	284

Tableau 8.13 Tableau des différences d'ajustement et valeur des tests de khi carré pour les modèles des SEM appliquées aux épreuves de CS

SEM	ΔX^2	Δdf	p	ΔRMSEA	ΔCFI	ΔTLI	ΔSRMR	ΔGFI	ΔAIC
CS Opérations – CS unidimensionnel	16,426	-1	< .00001	0,005	-0,018	-0,023	0,000	-0,021	18,426
CS Unités - CS unidimensionnel	-310,574	-2	< .00001	-0,084	0,137	0,165	-0,003	0,889	-306,574

8.9.3 Synthèse des résultats sur la conscience syntaxique

Les AFC basées sur les épreuves de conscience syntaxique ont montré que les modèles avec opérations n'étaient pas en mesure de représenter adéquatement les données. Le modèle unidimensionnel s'est montré également inadéquat. C'est le modèle bidimensionnel mettant en œuvre deux unités syntaxiques de petite et plus grande taille, soit les mots, représentant des classes grammaticales, et les syntagmes, occupant des fonctions au sein de phrases, qui s'est révélé le mieux ajusté aux données.

Du côté des SEM, le modèle bidimensionnel avec unités s'est avéré également ajusté aux données, et a permis de montrer que les variables latentes construites prédisent effectivement la variable endogène spécifiée représentant la compréhension en lecture.

8.10 Résumé des résultats et limites de la recherche

Les résultats pour la conscience morphologique sont assez limpides. Les AFC indiquent que les mesures de la CM sont mieux représentées par un construit unidimensionnel. Contrairement à nos hypothèses de départ, la scission des épreuves de CM en fonction de l'unité linguistique ou des opérations qui sont mises en œuvre ne permet pas de mieux représenter les différences interindividuelles déployées dans

nos données. Les SEM montrent que le modèle unidimensionnel est le mieux ajusté aux données, et montre un lien significatif entre la variable représentant la conscience morphologique et la variable dépendante Compréhension. Il y a une conscience morphologique qui semble fonctionner avec l'ensemble des épreuves.

Les résultats concernant la conscience syntaxique présentent également un portrait assez clair de la variation existant au sein des données, moyennant l'ajout de covariances entre certains termes d'erreur. Les AFC effectuées avec les données de CS confirment que les modèles qui fonctionnent le mieux sont ceux qui délimitent les unités linguistiques contenues dans les épreuves. Ces résultats font certainement écho aux ACP, de nature exploratoire, présentées à la section 8.5, qui laissaient envisager, dans les données, la présence d'une dimension sous-jacente d'ordre linguistique, prenant la forme de facteurs latents relatifs aux mots et aux syntagmes. Nos résultats ne nous permettent pas d'accepter l'hypothèse alternative de la présence de dimensions sous-jacentes d'ordre cognitif, représentées par des opérations d'addition et de suppression.

Les modèles travaillés par SEM pour les épreuves de CS permettent de conclure que les habiletés sous-jacentes hypothétisées, relatives aux unités – mots et syntagmes – contenues dans les épreuves, ont un impact différencié sur la compréhension en lecture. Autrement dit, la différence de traitement de la conscience syntaxique en

fonction de deux unités syntaxiques de tailles différentes, qui illustrent deux niveaux de la représentation syntaxique, prédit mieux la compréhension en lecture.

Un résultat persistant et transversal à nos deux construits et à nos deux modèles statistiques est la corrélation au-dessus de 1 pour les AFC et SEM pour les modèles impliquant les opérations d'addition et de suppression en particulier. Les corrélations entre variables latentes dans un modèle SEM ne peuvent pas être supérieures à 1 dans un cadre théorique valide, car elles reflètent des relations linéaires entre les variables latentes. Dans le contexte d'un modèle SEM, si des variables latentes sont fortement corrélées, cela peut entrainer des problèmes de multicolinéarité. Lorsque cela se produit, le modèle peut avoir du mal à estimer correctement les coefficients de régression ou les charges factorielles, ce qui peut conduire à des résultats inattendus, y compris des corrélations supérieures à 1. Il est possible que les variables latentes dans nos modèles déficients soient mesurées de manière très similaire et qu'elles se chevauchent, de sorte qu'elles sont statistiquement presque identiques. Cela pourrait expliquer les problèmes dans la spécification des modèles. Quoi qu'il en soit, nous ne sommes pas en mesure de conclure que les opérations cognitives d'addition et de suppression, telles que nous les avons conçues, circonscrivent des variables pertinentes pour décrire la variation dans les épreuves de CM et de CS.

The whole idea behind cognitive neuroscience is the once radical notion that the mysteries of human thought can be explained in much the same way as everything else in science – by reducing a complex phenomenon (cognition) into simpler components (the underlying biological mechanisms of the brain).

— Randall C. O'Reilly et Yuko Munakata, *Computational explorations in cognitive neuroscience*

CHAPITRE IX

DISCUSSION

Dans le chapitre VIII, nous avons fait état des résultats des analyses appliquées à la conscience morphologique, à la conscience syntaxique et à la compréhension en lecture. Nous nous consacrerons, dans le présent chapitre, à répondre à nos deux questions de recherche de nature psycholinguistique (9.1 et 9.2) au regard de ces résultats. Nous ferons état des biais potentiels ayant pu peser sur nos résultats (9.3) et nous positionnerons sur les constats à tirer de ces résultats au regard de la nature du traitement de l'écrit (9.4).

9.1 Réponses à la première question de recherche

Revenons d'abord sur notre première question de recherche, qui interroge la dimensionnalité de la conscience morphologique et de la conscience syntaxique au vu des épreuves utilisées pour les mesurer. Pour rappel, la question s'articule comme suit :

- *Les mesures traditionnelles des habiletés en conscience morphologique et en conscience syntaxique peuvent-elles se scinder en deux constituants, l'un linguistique, l'autre cognitif, prenant la forme de variables latentes ?*

En ce qui a trait à la conscience morphologique, les AFC indiquent que les tâches traditionnellement utilisées peuvent se scinder en deux variables latentes, représentant les deux types d'unités mises en œuvre : les morphèmes dérivationnels et les morphèmes flexionnels. Les résultats indiquent toutefois que les données, et donc la variation au sein des épreuves, sont mieux représentées par un construit unidimensionnel, représentant la CM dans sa globalité. Les différentes tâches de conscience morphologique représentent les éléments essentiels de la conscience morphologique ; celles-ci s'expliquent et se comprennent par un niveau général de conscience morphologique.

Ces résultats coincident avec ceux d'autres études ayant également conclu à l'unidimensionnalité de la CM (Muse, 2005; Spencer, Mercedes *et al.*, 2015; Tibi et Kirby, 2017). Ils tranchent, cependant, avec Goodwin, A. P. *et al.* (2017), qui avaient

conclu à la présence d'une conscience morphologique bidimensionnelle fonctionnant avec une dimension générale de connaissances de la structure morphologique des mots, et un autre, regroupant sept sous-dimensions se référant à l'application de connaissances morphologiques pour répondre aux spécificités des tâches. Tighe et Schatschneider (2015) avaient également conclu à une conscience morphologique bidimensionnelle, en fonction des tâches impliquant des mots et des pseudomots. Chez Tibi et Kirby (2017), le modèle différenciant les modalités orales et écrites permettait également de bien circonscrire la CM. Bourdages et Foucambert (2018) avaient conclu qu'un modèle séparant les épreuves en deux types d'opérations (ajout et suppression) était mieux ajusté aux données qu'un modèle unidimensionnel de la CM.

Dans une visée diagnostique, on pourrait arguer qu'une seule des 6 épreuves présentées suffirait à discriminer les élèves faibles des élèves fort·es en conscience morphologique pour la population que nous avons étudiée. En effet, nos résultats tendent à démontrer qu'il n'est pas la peine de déconstruire la conscience morphologique pour en exposer les dimensions sous-jacentes. Cependant, la combinaison de plusieurs épreuves pour un seul construit reste intéressante d'un point de vue docimologique, en permettant une mesure plus fiable et plus précise du construit.

Les résultats pour la conscience syntaxique permettent d'envisager sa bidimensionnalité à l'aune des unités linguistiques mises en œuvre au sein des épreuves, soit les mots, correspondant à une représentation des classes grammaticales, et les syntagmes, correspondant à une représentation de taille plus large, et de nature plus abstraite de la syntaxe : les fonctions grammaticales. Nos résultats montrent que le travail sur la conscience syntaxique doit se faire en décomposant les représentations de la langue qu'elle mobilise.

Nous avons donc constaté, tant pour les épreuves de CM que de CS, que la réussite aux épreuves requiert une habileté dans la reconnaissance du contenu linguistique qu'elles renferment. C'est donc cette habileté latente, de nature linguistique, à accomplir cette reconnaissance, et à en affiner la représentation afférente, qui contribue aux variables mesurées dans les HmL et qui influence les performances en compréhension en lecture.

Notre hypothèse quant à la contribution d'habiletés latentes de nature cognitive aux variables mesurées dans les épreuves traditionnellement utilisées pour la CM et la CS ne peut être acceptée. En effet, les deux construits présentent des patrons tout à fait analogues, dans lesquels les variables Ajout et Suppression présentent des corrélations trop élevées, ne permettant pas de bien circonscrire les habiletés latentes en question. De plus, les modèles comportant ces mêmes habiletés et incluant les épreuves avec matériau non linguistique n'ont pas permis un meilleur ajustement aux

données, renforçant l'idée que les variables latentes de nature cognitive ne sont pas à même de bien décrire la variation dans les données.

Cet état de choses nous permet de dresser au moins deux pistes d'explications. D'une part, nous pouvons postuler qu'effectivement, l'ajout et la suppression sont les deux faces d'une même pièce : ce sont deux opérations analogues, et donc hautement corrélées, appartenant à un même mécanisme cognitif. Leur nature pourrait les rendre indissociables, compliquant ainsi leur mesure, en tout cas au regard d'habiletés latentes différentes et donc dissociables. Il est possible également qu'elles soient hautement corrélées parce que complètement automatisées chez tous les sujets : leur mesure n'implique ainsi pas de variation notable. D'autre part, il est possible que nos épreuves, construites entièrement pour les besoins de la présente thèse, et n'ayant donc pas été étalonnées, ne permettent pas de rendre compte adéquatement de la variation dans les épreuves traditionnellement utilisées pour mesurer la CM et la CS.

9.2 Réponses à la deuxième question de recherche

La deuxième question de recherche s'articule comme suit :

- *Dans quelle mesure les variables latentes construites – d'ordres linguistique et cognitif – contribuent-elles à la compréhension en lecture?*

Les résultats concernant la conscience morphologique indiquent que le modèle unidimensionnel, qui entend la conscience morphologique comme un construit unifié et cohérent, est celui qui s'ajuste le mieux aux données. Ce modèle indique une contribution significative ($\beta = 0,76$) de la CM vers la compréhension en lecture. La conscience morphologique est réputée être maitrisée, donc automatisée, plus tôt que la conscience syntaxique dans le développement. Il faudrait vérifier si une contribution plus faible de la CM chez les élèves plus jeunes est à constater dans nos données. Les liens extrêmement étroits qu'entretiennent la conscience morphologique et le vocabulaire (Nagy *et al.*, 2006) peut contribuer à expliquer l'impact de la CM sur la compréhension.

Cette contribution de la conscience morphologique comme construit unidimensionnel, avec une influence sur la compréhension qui fait fi de nombreux facteurs liés à la diversité des tâches qui la définissent, suggère l'existence d'une conscience métalinguistique des unités morphologiques de la langue qui est profondément enracinée dans le système linguistique.

Les résultats pour la conscience syntaxique sont analogues, en ce sens que la dimensionnalité des tâches de CS révèle celle relative à la syntaxe et au système qui participe de la compréhension du langage écrit. L'habileté à prendre conscience des mots et des syntagmes a un impact sur la compréhension.

Nous avons mentionné que les deux variables latentes de nature linguistique présentent une corrélation de 0,74. Elles sont donc liées : un élève qui réussit à repérer et manipuler le mot réussira également à repérer et manipuler un syntagme. De plus, les contributions des variables linguistiques « mot » et « syntagme » à la compréhension dans le modèle structurel sont presque équivalentes. Contrairement à ce que nous aurions pu envisager, l'habileté à concevoir une unité linguistique supérieure, ou plus abstraite, par rapport à une unité linguistique plus basique, ici le syntagme par rapport au mot, ne semble pas mieux prédire l'habileté complexe qu'est la compréhension. Ceci tend à indiquer que ces deux composantes de l'analyse syntaxique de la langue irriguent, chacune à leur façon, la compréhension en lecture.

La CS est réputée faciliter le *parsing* syntaxique, qui permet la formation de représentations de niveau textuel, un mécanisme proposé par Perfetti *et al.* (2005). Ce redécoupage syntaxique des phrases en éléments plus maniables pourrait permettre la formation de ces représentations (Deacon et Kieffer, 2018). Il est possible que les CM et CS soient tellement intégrées (*embedded*) aux représentations lexicales qu'on ne puisse voir ce qui vient de l'un ou de l'autre.

D'un point de vue théorique, notre recherche confirme ainsi l'influence de la CM et de la CS sur la compréhension, en accord avec de nombreux modèles explicatifs de la compréhension en lecture.

La réflexion sur la dimensionnalité des construits, et les analyses statistiques permettant de les évaluer, augure le foisonnement de la réflexion sur d'autres facteurs représentant la diversité des tâches, en les envisageant sur des continuums. Nous pouvons citer en exemple l'opposition entre les morphèmes au sein de mots ou au sein de phrases (Bernstein *et al.*, 2020), ou encore l'opposition mots et pseudomots (Tighe et Schatschneider, 2015) pour nos épreuves de conscience morphologique. Ces deux dimensions permettent effectivement de différencier nos tâches, et ce, malgré tout le soin apporté au contrôle des variables de confusion à nos questions de recherche dans la constitution de nos épreuves. Nos épreuves de CM peuvent également être repensées en fonction des facteurs d'analyse et de contrôle postulés par Bialystok (2001a), que nous avons présentés dans la section 1.2.2. Cette reconceptualisation de nos outils de mesure permet d'arriver à des classements additionnels dont le résultat pour la CM figure dans le tableau 9.1.

Tableau 9.1 Aperçu des tâches de conscience morphologique utilisées et de leur classement en fonction de quatre dimensions

		Dimensions		
Épreuves	**Unité**	**Types d'analyse**	**Niveaux de contrôle**	**Opération**
1. Paires de mots	Mots dérivés et pseudodérivés avec suffixes dérivationnels	**Implicite**	**Bas**	Suppression
2. Extraction de la base	Suffixes dérivationnels	**Explicite**	**Haut**	Suppression

3. Dérivation	Morphèmes dérivationnels et flexionnels	**Explicite**	**Bas**	Ajout
4. Décomposition	Morphèmes dérivationnels et flexionnels	**Explicite**	**Bas**	Suppression
5. Choix de pseudomots suffixés D	Morphèmes dérivationnels	**Implicite**	**Haut**	Ajout
6. Choix de pseudomots suffixés F	Morphèmes flexionnels	**Implicite**	**Haut**	Ajout

Cette réflexion sur les dimensions sous-jacentes additionnelles s'applique également, bien entendu, à nos mesures de conscience syntaxique. Le tableau 9.2 fait état de cette remise en question.

Tableau 9.2 Aperçu des tâches de conscience syntaxique utilisées et de leur classement en fonction de quatre dimensions

	Dimensions			
Épreuves	**Unités**	**Types d'analyse**	**Niveaux de contrôle**	**Opération**
Identification de l'erreur dans phrase agrammaticale M	Mot	**Implicite**	**Bas**	Suppression
Correction de l'erreur dans phrase agrammaticale M	Mot	**Implicite**	**Bas**	Ajout
Identification phrase grammaticale, réplication 1 M	Mot	**Implicite**	**Haut**	Suppression
Correction phrase grammaticale, réplication 1 M	Mot	**Implicite**	**Haut**	Ajout
Identification phrase grammaticale, réplication 2 M	Mot	**Explicite**	**Haut**	Suppression
Correction phrase grammaticale, réplication 2 M	Mot	**Explicite**	**Haut**	Ajout

Identification de l'erreur dans phrase agrammaticale S	Syntagme	**Implicite**	**Bas**	Suppression
Correction de l'erreur dans phrase agrammaticale S	Syntagme	**Implicite**	**Bas**	Ajout
Identification phrase grammaticale, réplication 1 S	Syntagme	**Implicite**	**Haut**	Suppression
Correction phrase grammaticale, réplication 1 S	Syntagme	**Implicite**	**Haut**	Ajout
Identification phrase grammaticale, réplication 2 S	Syntagme	**Explicite**	**Haut**	Suppression
Correction phrase grammaticale, réplication 2 S	Syntagme	**Explicite**	**Haut**	Ajout

L'inclusion de ces dimensions dans des modèles SEM additionnels permettrait d'évaluer si ces modèles peuvent fournir des résultats plus cohérents et explicatifs. De plus, l'exploration d'autres modèles théoriques ou d'autres méthodes d'analyse, notamment les SEM multiniveaux, pourrait également contribuer à une meilleure compréhension des résultats observés et à une interprétation plus solide des relations entre les variables étudiées.

9.3 Biais potentiels et pistes d'amélioration

Plusieurs pistes d'explication peuvent être avancées pour comprendre les résultats mitigés obtenus pour plusieurs des hypothèses mises à l'épreuve dans notre étude. Tout d'abord, le caractère imbriqué des données sur la CS n'a pas été correctement pris en compte sur le plan statistique, ce qui peut avoir introduit des biais dans l'analyse. En effet, pour réussir la réplication 2, il faut avoir réussi la réplication 1, pour réussir la réplication 1, il faut avoir réussi la correction, et pour réussir la correction, il

faut au préalable avoir identifié l'erreur (identification). Ces « épreuves gigogne » posent problème, parce que la mesure d'une tâche implique et reprend la mesure de l'autre. D'ailleurs, les indices de modification dans les épreuves de CS recommandaient assez systématiquement de faire covarier les termes d'erreur chaque fois des variables fortement imbriquées, indiquant un réel souci méthodologique. Ce souci aurait pu être pris en compte par des analyses SEM hiérarchiques, complexifiant le traitement des données. Il aurait été important de prendre en compte cette structure complexe des données lors de l'analyse statistique afin d'obtenir des résultats fiables.

De plus, signalons que les épreuves CM et CS ne sont pas égales sur le plan des opérations. Celles de CS ont été déconstruites de façon plus fine que celles de CM. En effet, la tâche de décomposition (CM, en 1), par exemple, implique une opération de suppression, selon notre cadre conceptuel. En parallèle, les items de CS (en 2) regroupant l'identification de l'erreur n'impliquent à proprement parler qu'une opération de suppression.

(1) Épreuves de CM

Décomposition (suppression)
Ex : *courageux. Mario a beaucoup de* _____. [courage]
Ex. *fatigante*. *Mon chien est un peu* _____.[fatigant]

(2) Épreuve de CS

Identification de l'erreur (suppression)
Ex. lave la vaisselle le papa

Ex. mon chien va par le vétérinaire

L'épreuve de décomposition implique en quelque sorte la suppression du morphème, puis l'ajout de *courage* en fin de phrase, tandis que l'identification de l'erreur n'implique qu'une suppression, purement mentale. Notre délimitation des opérations minimales de suppression et d'ajout n'est donc pas optimale. Peut-être faudrait-il revenir à des opérations à une granularité plus élevée : remplacement, déplacement, etc.

De plus, les épreuves de compréhension utilisées dans l'étude mesurent un construit qui n'est pas complètement cohérent, comme en témoignent les indices de fiabilité qui se rapprochent, dans certains modèles, des limites recommandées (Yanuar *et al.*, 2015; Zinbarg *et al.*, 2005). Nos mesures de la compréhension en lecture sont certes plurielles, mais aussi très larges. Nous avons voulu évaluer la construction d'un modèle mental, soit une compréhension profonde, en voulant éviter des épreuves standardisées de type passages de texte – comme les épreuves de compréhension à la Woodcock (Woodcock, 1987; Woodcock *et al.*, 2001).

Notre mesure de la compréhension est pourtant critiquable. Sur le plan de la réussite brute au test de compréhension sur texte long, les performances aux questions ayant trait à des éléments explicitement mentionnés dans le texte (réponses A) se démarquaient assez peu des éléments censés mesurer une compréhension sophistiquée (réponses C). En effet, les élèves ont démontré des réussites similaires pour ces deux types de réponses, alors qu'on aurait pu s'attendre à des résultats plus

bas pour les réponses relevant de la construction d'un modèle mental cohérent. Bien sûr, il est notoirement difficile d'évaluer la construction d'un modèle mental, celui-ci étant par définition complètement personnel, unique. Un texte peut être interprété, vécu, compris complètement différemment par deux individus qu'à priori rien n'oppose. La cohérence imparfaite des épreuves peut donc avoir contribué à l'obtention de résultats mitigés, en particulier pour les modèles SEM impliquant les épreuves de CS, mettant en évidence la nécessité de revoir la conception et la formulation de ces épreuves pour une mesure plus précise et fiable du construit Compréhension.

Une autre explication possible réside dans le nombre insuffisant d'items dans les épreuves matricielles non linguistiques, avec seulement 12 items pour la suppression et 12 items pour l'addition. Ce faible nombre d'items peut avoir eu un impact sur la fiabilité de ces épreuves. En effet, la fiabilité d'un test tend à augmenter avec sa longueur, reflétant ainsi sa capacité à mesurer de manière précise le phénomène étudié. Cette limitation en ce qui a trait au nombre d'items pourrait expliquer pourquoi l'ajout de ces variables matricielles dans les modèles de mesure impliquant des opérations n'a pas amélioré les modèles d'Analyse Factorielle Confirmatoire (AFC) ou de Modélisation par Équation Structurelle (SEM).

L'évaluation de la possible médiation de la lecture de mots ou phrases dans la relation entre les habiletés métalinguistiques et la compréhension n'a pas été soumise à

l'expérimentation. D'importantes variables contrôle n'ont pas été incluses : conscience phonologique, mémoire de travail, vocabulaire. Les tâches visuospatiales peuvent être comprises comme des instruments mesurant le raisonnement non verbal. Peut-être aurait-il été judicieux d'inclure ces variables cruciales comme contrôles dans les modèles.

De plus, nous pourrions chercher à diviser les élèves en fonction de leur niveau de compréhension (à l'instar de Tong et al 2014) ou de leur niveau scolaire, afin de vérifier comment les deux construits régissent la compréhension chez les mauvais et bons compreneurs.

9.4 Positionnement plus général sur le traitement de l'écrit

Les thèmes et débats qui ont été abordés dans ce travail s'inscrivent dans une série de questions plus larges, et notamment celle, plus métaphysique, du fonctionnement de l'activité mentale, dont la conception par opposition aux sciences de la nature est tributaire du dualisme cartésien. Ainsi, les sciences de l'esprit, représentées aujourd'hui par les sciences cognitives, ont pour mandat d'étudier les manifestations de la psyché humaine et de l'activité mentale – dans la mesure où nous admettons la relation intrinsèque entre l'esprit et le cerveau humain (Gallina, 2011).

La psychologie cognitive et la psycholinguistique cherchent à répondre à des questions fondamentales, notamment celle du niveau d'analyse de ce qui est traité

par le cerveau (Fodor, 1981; Fuchs, 2017; Gallina, 2011). Le représentationnalisme et les représentations ont donc occupé une place centrale dans l'étude des sciences cognitives, et ce, tant en cognitivisme computationnel classique que du point de vue connexionniste. Il s'agit d'une thèse épistémologique, d'abord, selon laquelle l'esprit fonctionne essentiellement à l'aide des représentations et d'opérations effectuées sur ces représentations afin de produire des comportements intelligents, raisonnés, pertinents, flexibles et appropriés, en réaction à des situations, à un environnement. Les représentations, en elles-mêmes, peuvent être caractérisées comme des éléments discrets, isolables. Celles-ci sont retenues comme éléments d'analyse de l'esprit parce qu'elles semblent pertinentes ; pas que sur le plan conceptuel, mais aussi sur le plan biologique, c'est-à-dire en ce qui a trait à la réalité physique du cerveau[80].

Les éléments linguistiques manipulés dans le cadre d'épreuves de conscience métalinguistique sont en eux-mêmes des représentations, et les manipulations qui sont effectuées sur elles correspondent à des opérations, au sens cognitif du terme. La théorie des modèles mentaux constitue aussi une théorie des représentations. En effet, les représentations font partie intégrante de la théorie, en s'affinant au fur et à mesure que l'acte de compréhension se réalise. Les bris dans la compréhension

[80] Les approches connexionnistes de la cognition utilisent dans leur architecture des réseaux de nœuds et de liens dont l'organisation et le fonctionnement présentent une certaine isomorphie avec les systèmes neuronaux physiologiques. En ceci, les modèles connexionnistes sont certainement plus proches de la réalité physique du cerveau que les modèles computationnels, qui eux, mobilisent de manière importante les représentations, la computation et la modularité.

peuvent aussi être imputables à des pertes dans la qualité représentationnelle ou opératoire (Karimi et Ferreira, 2016). Ainsi, il est clair que les construits étudiés dans le cadre de cette thèse découlent d'une approche représentationnelle et computationnelle de la cognition et des états mentaux. Nos résultats semblent montrer que ce sont les habiletés à manipuler des concepts ou symboles abstraits, comme ceux de « morphème », de « mots » et de « syntagmes », même sans les nommer, qui sont à la source du développement cognitif de l'enfant (Bruner, 1973, 1983).

Nous avons constamment mis en parallèle, dans le cadre de cette thèse, les habiletés en conscience morphologique avec les habiletés en conscience syntaxique, postulant de facto que les représentations morphologiques et syntaxiques sont isomorphes. Nos résultats permettent de conclure également à une relative isomorphie entre les représentations grammaticales et les représentations mentales prenant forme dans le courant de la lecture et dont la finalité constitue le modèle de situation.

Nous n'avons pas été en mesure de conclure que des opérations au sein des épreuves métalinguistiques pouvaient expliquer des processus analogues en compréhension. Pourtant, la grande majorité des modèles de la compréhension inclut l'hypothèse d'une forme de traitement automatique des idées et représentations. Cette position est également vieille, comme en témoigne cet extrait de Thorndike (1917) :

> *Understanding a paragraph is like solving a problem in mathematics. It consists in selecting the right elements of the situation and putting them together in the right relations, and also with the right amount of weight or influence or force for each. The mind is assailed as it were by every word in the paragraph. It must select, repress, soften, emphasize, correlate and organize, all under the influence of the right mental set or purpose or demand* [81].

Les modèles de la compréhension de Gernsbacher, de Van Dijk et Kintsh et la plupart des autres incluent cette idée que des processus cognitifs qui ne sont pas spécifiquement langagiers sont mis en place dans le cadre de la compréhension de texte. Un grand nombre des mécanismes impliqués dans la compréhension en lecture sont probablement généraux. Il est possible que la séparation de nos épreuves en fonction de deux opérations minimales ait été trop audacieuse. Un niveau de granularité plus grand chez ces mécanismes cognitifs généraux aurait peut-être permis leur mesure sur le plan psycholinguistique.

---

[81] « Comprendre un paragraphe, c'est comme résoudre un problème de mathématiques. Il s'agit de sélectionner les bons éléments de la situation et de leur assigner les bonnes relations, ainsi que le poids, l'influence ou la force nécessaires pour chacun d'eux. L'esprit est pour ainsi dire assailli par chaque mot du paragraphe. Il doit sélectionner, réprimer, adoucir, souligner, corréler et organiser, le tout sous l'influence d'un répertoire mental, d'un objectif ou d'une demande appropriés » [notre traduction].

Et ainsi la plus grande distinction de ce qui se passe dans notre esprit, est de dire qu'on y peut considérer l'objet de notre pensée, et la forme ou la manière de notre pensée, dont la principale est le jugement : mais on y doit encore rapporter les conjonctions, disjonctions, et autres semblables opérations de notre esprit, et tous les autres mouvemens de notre âme, comme les desirs, le commandement, l'interrogation, etc.

— Antoine Arnauld et Claude Lancelot,
Grammaire générale et raisonnée de Port-Royal

CHAPITRE X

PERSPECTIVES DIDACTIQUES SUR LES TÂCHES MÉTASYNTAXIQUES

Nous avons, dans les chapitres précédents, procédé aux revues théorique et méthodologique de trois grands construits étudiés abondamment en psychologie cognitive : la conscience morphologique, la conscience syntaxique et la compréhension en lecture. Ces construits ont été soumis à l'expérience, et nous avons fait état des résultats de cette expérimentation. Le présent chapitre porte sur l'opérationnalisation et le traitement de ces habiletés dans les travaux en didactique et en psycholinguistique, dans une perspective méthodologique et comparatiste. Les résultats sont considérés à l'interface des sciences cognitives et de la didactique du français.

10.1 Présentation de la démarche de transposition didactique

Les difficultés en grammaire qu'éprouvent les élèves du primaire se manifestent notamment par une incapacité à reconnaitre et à segmenter des phrases, par une maitrise rudimentaire des règles syntaxiques – celles portant sur l'accord, par exemple –, par la difficulté à opérer des manipulations syntaxiques (Elalouf et Péret, 2009) et par la mauvaise compréhension des groupes constitutifs de la phrase (Gagnon et Péret, 2016)[82]. La réponse des sciences de l'Éducation, en général, et de la didactique de la grammaire en particulier, pour répondre à ces problèmes chez les élèves, est claire : c'est en nommant et en amenant des concepts et des structures à la conscience que l'élève pourra progresser (De Pietro, 2009; Leontiev, 1984). Pour Leontiev, un contenu ou un processus doit être envisagé de manière consciente dans le cadre d'une activité pour redevenir inconscient et utilisable dans une autre visée (Foucambert, 2003) :

> « les formes grammaticales […], que l'enfant utilise si parfaitement en tant que moyens de communication, d'expression verbale, ne peuvent cependant être contrôlées par la conscience; pour cela , il faut qu'elles deviennent d'abord l'objet spécial d'un rapport de l'enfant – l'objet de son action orientée vers un but » (Leontiev, 1875/1984, p. 298).

L'activité, pour Leontiev, se conçoit dans sa dimension sociale et historique, et forme un système ayant sa structure, ses passages, ses transformations internes. Cette

[82] Ces deux objets ont été discutés également dans Bourdages *et al.* (2021). Certaines parties du texte ont été reprises intégralement dans le présent chapitre.

définition permet de penser des ponts entre la tâche et l'activité, passage par lequel des éléments non conscients sont amenés à la conscience par l'exercisation et par le travail. L'idée générale est de rendre la connaissance explicite au moment de la phase d'apprentissage puis, de cette phase de non-conscience, à cette phase d'explicitation, en arriver à une phase où certains apprentissages sont automatisés.

Sur le plan de l'enseignement du fonctionnement de la langue, les finalités de l'école sur le plan de l'enseignement de la grammaire sont doubles : d'une part, construire la langue comme objet, et d'autre part, travailler la langue comme outil, en vue de communiquer, tant en production qu'en réception, à l'oral comme à l'écrit (Bronckart, 2004; Bulea Bronckart et Gagnon, 2017). L'élève est ainsi amené·e à mobiliser des habiletés métalinguistiques par des activités d'observation et de réflexion dans des contextes écrits significatifs (Reuter *et al.*, 2013). Nous l'avons vu, ces habiletés sont associées et prédisent des habiletés de plus haut niveau, comme la lecture et l'écriture, et sont réputées notamment permettre aux lectrices et lecteurs d'établir des prédictions, au cours de la lecture, sur les mots et les groupes qui suivront (Armand, 2005; Bratlie *et al.*, 2022; Simard, D. *et al.*, 2012b).

Les questions fondamentales autour de l'enseignement de la grammaire, allant de la détermination des objets à travailler à ses finalités (Bulea Bronckart, 2015; Bulea Bronckart et Gagnon, 2017; Chartrand, 2012b), apparaissent en filigrane dans le présent chapitre. La didactique de la grammaire tient pour acquis que l'explicitation

méta d'un savoir et la compétence à utiliser tel ou tel phénomène grammatical sont liées, l'une conduisant naturellement, automatiquement, l'autre. En effet, il semble raisonnable de penser que l'étude de la grammaire en classe entraine, logiquement, le développement des habiletés métasyntaxiques chez les élèves (Reuter *et al.*, 2013). Ces deux construits, les compétences grammaticales et les habiletés métasyntaxiques, sont d'ailleurs associés à la maitrise du code écrit (Dolz et Simard, 2009) dans leurs courants épistémologiques respectifs, la didactique du français et la psycholinguistique, et considérés comme étant d'une importance capitale pour la réussite scolaire. En effet, « la tradition scolaire a eu tendance à accorder à la grammaire explicite une grande importance en s'appuyant sur l'hypothèse non parfaitement étayée que les connaissances explicites concourent de façon majeure à la maitrise de la langue » (Simard, C. *et al.*, 2010). La question n'est pas ici de déterminer avec certitude que les connaissances grammaticales explicites, travaillées et retravaillées en salle de classe, ont un impact sur les connaissances implicites des élèves (De Pietro, 2009). Nous considérons cette probabilité pour acquise : l'explicitation et la distanciation par rapport à l'objet grammatical permettent de développer des connaissances grammaticales (Chiss & David, 2018). Les résultats exposés dans notre discussion, au chapitre IX, permettent à tout le moins d'attester, à nouveau, du lien prédictif qu'opèrent les habiletés grammaticales – comprises comme la conscience morphologique et la conscience syntaxique – sur les performances en lecture chez des élèves du primaire. Nous tentons dans le présent chapitre de penser le rapport entre habiletés métasyntaxiques et ce que l'on nomme

« raisonnement grammatical » (Marmy Cusin *et al*., 2022; Nadeau et Fisher, 2011), en essayant de proposer des activités scolaires permettant de susciter la réflexion sur le langage.

L'idée que nous proposons dans ce chapitre se situe définitivement entre les pôles « élève » et « savoirs », qui constituent l'axe psychologique du triangle didactique (Chevallard, 1985). Nous souhaitons repenser les exercices et les activités en classe autour du fonctionnement de la langue à la lumière des épreuves psycholinguistiques. La question de recherche à laquelle nous nous attèlerons, de nature didactique, est la suivante :

- *Comment peut-on opérer le passage entre tâches métalinguistiques et activités grammaticales dans la classe du cycle 2 au primaire?*

Nous cherchons ainsi à établir si une transposition didactique des épreuves psycholinguistiques aux activités grammaticales est possible et, si tel est le cas, si elle est souhaitable et si elle peut être reconceptualisée. Nous tenterons de faire valoir que les tâches métalinguistiques peuvent être porteuses pour l'étude de la langue dans la classe de primaire et d'opérer les transformations pertinentes pour favoriser un passage harmonieux entre tâches métalinguistiques et activités grammaticales.

10.2 Évaluation et enseignement de la grammaire

Nous traitons dans cette section des manières d'aborder la mesure ou l'évaluation des habiletés syntaxiques et des formes de raisonnement grammatical dans deux domaines qui s'intéressent aux habiletés métalinguistiques et à leur impact dans le développement linguistique des élèves : la psycholinguistique et la didactique de la grammaire. Ceci implique, d'une part, de faire état des objectifs de recherche différenciés des deux disciplines, prenant pour ancrages leurs postures épistémologiques respectives, et d'autre part, de décrire les tâches ou activités prévues pour mesurer/développer la conscience syntaxique. La didactique du français est ici entrevue comme une « discipline d'action » (Bronckart et Chiss, 1990, p. 19) qui prend appui sur les disciplines de référence, la psycholinguistique au premier chef. Nous partageons ainsi la thèse forte formulée par Bronckart et Chiss (1990 p. 34) selon laquelle « les rapports qu'entretiennent descriptions linguistiques et analyses psycholinguistiques avec les préoccupations explicitement didactiques […] constituent des enjeux essentiels de réflexion au plan méthodologique et théorique ».

10.2.1 Enseignement de la grammaire

Les premiers contacts avec le fonctionnement de la langue ou de la grammaire, comme discipline scolaire, du ou de la jeune élève se font par l'entremise des interactions à l'oral avec l'enseignant·e. L'élève est en effet amené·e, dès son entrée à l'école, à reformuler ses énoncés de manière à ce que ceux-ci soient

syntaxiquement corrects (CIIP, 2010-2016). Par la suite, grâce à des dispositifs d'apprentissage du code écrit, tels que la dictée à l'adulte, l'élève se familiarise progressivement avec la grammaire. Il comprend que le texte n'est pas un amas de mots, mais un ensemble de phrases (délimitées d'un point de vue graphique). La notion de phrase commence à faire l'objet d'un enseignement systématique autour de 6-7 ans dans plusieurs pays francophones (CIIP, 2010-2016; MELS, 2009). L'introduction de l'élève à la grammaire de la phrase intègre le repérage des constituants obligatoires et facultatifs de la phrase, l'identification des classes grammaticales du nom et du verbe, ce qui amène progressivement la reconnaissance des constructions du groupe du nom et du groupe du verbe (CIIP, 2010-2016; MELS, 2009). La notion de groupe permet à l'élève de percevoir les régularités et récurrences dans l'organisation des mots de la phrase (Nadeau, 2017). La compréhension de la construction et des caractéristiques de la phrase syntaxique – la phrase de base est constituée d'un sujet, d'un prédicat et facultativement d'un complément de phrases (Chartrand, 2012a) – conduit l'élève à reconnaitre une phrase transformée ainsi qu'à repérer les types et formes de phrases (Simard, C. et Chartrand, 2011). À la fin du premier cycle du primaire, l'élève sait identifier les fonctions de sujet, de prédicat, de complément de phrase. Puis, dès le deuxième cycle, le travail des fonctions est instrumenté à l'aide de manipulations syntaxiques telles que la pronominalisation, l'ajout, l'encadrement (ou emphase), le déplacement et l'effacement (Chartrand, 2013).

10.2.2 Activités syntaxiques et évaluation

En conformité avec l'idée de la double finalité de l'enseignement grammatical (Bronckart, 2004), l'école doit stimuler le développement des habiletés métasyntaxiques par des activités d'observation et de réflexion qui portent sur des contextes signifiants, réels, afin de favoriser la construction de la langue comme système et permettre à l'élève de communiquer, tant en production qu'en réception. Ultimement, les manifestions d'observables métalinguistiques en classe, comme l'utilisation de la terminologie grammaticale et les manipulations syntaxiques, permettent de dresser le portrait des réflexions métalinguistiques que peuvent déployer les élèves (Boivin, 2014). De plus, afin d'améliorer la syntaxe des élèves et de favoriser son transfert en production textuelle, de nombreux·euses didacticien·nes « proposent de rapprocher les activités grammaticales de la situation d'écriture et d'y intégrer des principes clés inspirés du socioconstructivisme, soit la verbalisation des représentations grammaticales, les démarches inductives (par lesquelles sont dégagées des régularités à la suite de l'observation d'exemples d'un phénomène grammatical), le traitement de l'écrit sous l'angle de la résolution de problème, la discussion entre pairs, etc. » (Quevillon Lacasse *et al.*, 2018, p. 1).

Ces situations de verbalisations grammaticales visent l'utilisation et la maitrise du métalangage : la précision des termes employés par les élèves se conçoit comme un progrès ; cet usage contribue au transfert des apprentissages en écriture, car il conduit l'élève à intérioriser la manière de tenir un raisonnement grammatical complet qui

inclut les manipulations (Fisher, Huneault & Nadeau, 2015). L'enseignement est explicite, et l'élève est actif : celui ou celle-ci est amené·e à adopter une position de chercheur·e vis-à-vis de l'objet linguistique.

Les activités grammaticales en classe, qui peuvent prendre appui sur des corpus, par exemple, sont rattachées à des contextes de production. Les dispositifs didactiques sont pensés à l'intérieur d'un processus spiralaire au sein de la scolarité obligatoire : des collections organisées d'activités se succèdent sous forme de parcours ayant un début, une fin et un objectif précis. L'activité didactique porte sur un objet spécifique du savoir : elle est liée à une consigne, engendre une forme sociale de travail qui lui est propre, et bénéficie d'un support matériel qui lui est lié (Dolz et Toulou, 2008). Ces approches dites contextualisées, dans lesquelles les phénomènes linguistiques sont travaillés à travers les textes, sont promues par la didactique de la grammaire (Bulea Bronckart et Garcia-Debanc, 2021). Les pratiques actuelles montrent des degrés divers d'intégration de l'enseignement grammatical rénové : les approches inductives, articulant le travail de la langue et le travail du texte, entrent en conflit avec le découpage traditionnel de la leçon de grammaire et requièrent des outils encore peu maitrisés par les professionnels (Elalouf *et al.*, 2017; Nadeau, 2017). À l'instar des tâches classiques utilisées pour évaluer les habiletés métalinguistiques (Simard, D. *et al.*, 2011), que nous avons reconceptualisées dans les chapitres II et III à l'aune des axes imaginés par Bialystok (2001a), les activités grammaticales réalisées en classe peuvent aussi être projetées sur le modèle. La figure 10.1 correspond à une tentative de classement de quelques-unes des activités grammaticales répertoriées en fonction

des axes de contrôle et d'analyse des éléments linguistiques. Ces axes sont des outils pour penser : ils permettent de guider la réflexion sur les épreuves en donnant une large place au débat.

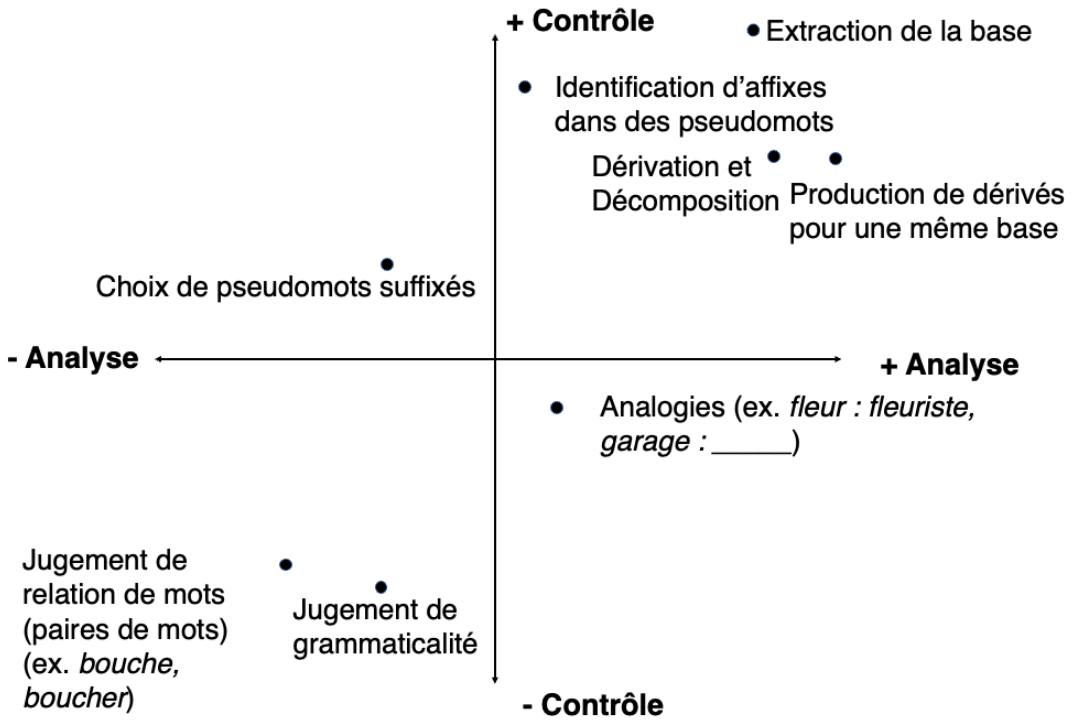


Figure 10.1 Modèle de Bialystok (2001a) avec projection d'activités d'enseignement grammatical issues de Bulea Bronckart, Marmy Cusin & Panchout-Dubois (2017) et Lepoire Duc & Sautot (2010)

Pour rappel, les activités présentées dans le quadrant en bas à gauche de la figure 3 sont plutôt de type linguistique, alors que celles qui sont projetées dans le quadrant en haut à droite sont réputées mobiliser plus de contrôle et d'analyse, et donc être de nature plus métalinguistique.

La didactique de la grammaire tend à promouvoir la contextualisation des règles grammaticales et l'instauration des « démarches à caractère inductif, accordant une large place à l'activité de l'élève : observer des énoncés, les manipuler, en dégager des régularités, enfin conceptualiser » (Bulea Bronckart *et al.*, 2017, p. 131). Nous proposons dans la figure 10.1 un classement des différentes activités grammaticales sur les deux axes conçus par Bialystok (2001a, b). Cette représentation permet de visualiser le fait que plusieurs des tâches réalisées en salle de classe demandent des habiletés élevées de contrôle et d'analyse. Les activités grammaticales sont de nature complexes, c'est-à-dire qu'elles sont souvent caractérisées par plusieurs étapes menant ultimement à la verbalisation, et mobilisent probablement chez les élèves des habiletés de contrôle et d'analyse souvent élevées : ainsi, la réponse à une seule consigne est susceptible d'impliquer des processus plus implicites – comme la détection et la correction – suivis de processus explicites (analyse, manipulations syntaxiques, identification, classement, etc.). Ces activités d'enseignement mobilisent tour à tour des capacités de découverte, d'exploration, de construction, de conceptualisation, de classification, de mémorisation et d'argumentation (ou justification). La terminologie grammaticale est ainsi utilisée pour laisser des traces de l'accès au métalangage qui a lieu au cours de la phase d'institutionnalisation (Nadeau et Fisher, 2006), notamment par la verbalisation de stratégies permettant d'identifier des classes et les fonctions grammaticales.

10.3 Comparaison entre les tâches métasyntaxiques et les activités grammaticales

Il faut ici rappeler l'intérêt de penser des ponts entre les deux disciplines, puis leurs spécificités. Pour comprendre et pour éventuellement dépasser ces différences terminologiques et, de fait, épistémologiques entre psycholinguistique et didactique sur le plan du traitement des habiletés métasyntaxiques, un zoom méthodologique est nécessaire. Nous proposons d'entreprendre cette démarche en réfléchissant à ce qui rapproche et oppose les tâches psycholinguistiques et les activités grammaticales en classe.

10.3.1 Enseignement de la grammaire

Les deux disciplines partagent d'abord un problème ou une difficulté proprement méthodologique : celle de pouvoir différencier ce qui appartient au processus de ce qui appartient au résultat, un constat partagé par Sander (2016, p.464) : « […] pas plus qu'une photographie ne contient l'histoire de vie des personnages et des lieux qu'elle figure, la trace de l'aboutissement d'un processus ne permet son retracement. »

Il y a corollairement cet intérêt partagé pour les opérations mentales que l'élève met en place pour arriver à répondre aux consignes. Le découpage de ces opérations, le type d'habiletés qui seront mobilisées pendant la réalisation des activités, les raisonnements grammaticaux auxquels l'élève procèdera pour arriver à un résultat

revêtent également un intérêt certain tant pour la didactique que la psycholinguistique. Le modèle de Bialystok (2001a) est nécessaire afin de procéder à une analyse systématique des mesures utilisées dans la littérature sur les habiletés métalinguistiques, puisqu'il fait une distinction claire entre les habiletés qui dépendent de la connaissance et de l'analyse formelles de la langue et celles qui dépendent du contrôle attentionnel requis pour isoler forme et sens.

En classe, la grammaire, intrinsèquement liée à la maitrise du code écrit (Dolz et Simard, 2009), peut être travaillée à l'aide d'exercices contextualisés, mais aussi d'autres dont l'analyse porte sur l'unité « phrase » et qui ressemblent aux tâches utilisées en psycholinguistique. Plusieurs tâches sont ainsi communes aux deux disciplines, alors que d'autres sont spécifiques à la salle de classe. Ces dernières peuvent être faites en groupe, et incluent souvent plusieurs types et séries de manipulations : elles sont plus complexes, en ce sens, ou moins pures, que les tâches en psycholinguistique. Le tableau 10.1 permet d'avoir un aperçu de cet état de choses. Les tâches strictement psycholinguistiques sont en gras. Toutes deux mettent en œuvre l'erreur.

Tableau 10.1 Recension des tâches utilisées dans la recherche psycholinguistique pour mesurer les habiletés métasyntaxiques et des activités grammaticales réalisées en salle de classe

	Tâches en psycholinguistique	**Activités en salle de classe**
Jugement de grammaticalité	X	X
Localisation de l'erreur	X	X
Correction de l'erreur	X	X
Répétition de phrases agrammaticales	X	
Test de closure (*cloze task*)	X	X
Réplication de l'erreur (ou production analogique de phrases non grammaticales)	X	
	X	X
Explication de l'erreur	X	X
Réordonnancement de phrases (*word ordering*)		X
Combinaisons de phrases		X
Classification de groupes de mots avec ou sans critères		X
Encadrement (avec « c'est… qui », par exemple)	X	X
Identification de classes grammaticales		X
Suppression, ajout, déplacement des classes grammaticales		X
Verbalisation de stratégies pour identifier les classes grammaticales		X
Tri de mots (classer des groupes de mots et les comparer)		X
Justification de transformations		X
Réécriture d'un texte, avec transformations syntaxiques		

Note. Cette recension de tâches provient notamment de Besse, A.-S. *et al.* (2010), Bulea Bronckart *et al.* (2017), Lesaux et Siegel (2003), Quevillon Lacasse *et al.* (2018) et Simard, D. *et al.* (2017).

10.3.2 Activités syntaxiques et évaluation

Les deux disciplines présentent des différences sur le plan de leur traitement méthodologique de la conscience syntaxique. L'école doit stimuler les habiletés métalinguistiques par des activités d'observation et de réflexion dans des contextes écrits significatifs. Ainsi, contrairement à la recherche en psycholinguistique, qui se concentre sur les unités relatives à la structure de la phrase, tenant compte de contraintes expérimentales comme la rapidité et la fiabilité des mesures, en utilisant des items portant sur l'unité phrase ou même sur le mot, une approche didactique de la syntaxe, dans le cadre d'un enseignement dit renouvelé de la grammaire, requiert la prise en compte du texte comme unité (Arseneau, 2016). Comme mentionné précédemment, les leçons de grammaire en classe vont aussi impliquer des opérations ou actions de plusieurs ordres, certaines complexes, d'autres beaucoup moins. En effet, certaines activités faites en classe, dans les moins complexes, mobilisent des connaissances épilinguistiques dans des exercices faciles réalisés rapidement par l'élève, sans que la notion ait été traitée, comprise, opérationnalisée (Audion, 2016). Mais il est rare de proposer aux élèves des activités qu'on pourrait qualifier de « pures », c'est-à-dire des activités dans lesquelles on postule qu'un seul type d'habileté sera mobilisé, alors que la psycholinguistique, dont les visées sont descriptives et psychométriques, privilégie ce type d'activités.

Souvent aussi, l'utilisation de la terminologie grammaticale survient au terme des activités grammaticales, dans les moments d'institutionnalisation plus particulièrement

(Chartrand et De Pietro, 2012; Chiss et David, 2011). Par ce travail de verbalisation, l'élève est accompagné dans son processus réflexif jusqu'à une zone de développement potentiel (ZPD) (Vygotsky, 1934/1997), et l'enseignant·e est en mesure de juger de ce qui se passe chez l'élève et de la présence ou non d'un raisonnement grammatical. L'utilisation du métalangage peut aussi favoriser le développement des habiletés métalinguistiques en général (Arseneau, 2016), particulièrement celles ayant trait à l'analyse des connaissances linguistiques, notamment pour la réalisation des accords grammaticaux (Nadeau et Fisher, 2011).

Le ou la psycholinguiste arrive à obtenir des informations sur les habiletés métalinguistiques des sujets et sur la variabilité entre eux et elles en cloisonnant à priori les tâches, et par l'entremise d'une mesure chiffrée des performances de chacune et chacun à ces épreuves. Ainsi, la prise d'informations sur les habiletés métasyntaxiques de l'élève est effectuée différemment dans les deux disciplines. Celles-ci peuvent aussi être contrastées sur le plan de leur intérêt pour la progression de ces habiletés chez l'élève. Alors que la psycholinguistique effectue un travail théorique sur ce plan, par l'entremise d'études longitudinales ou transversales basées sur des données prises de façon ponctuelle, la didactique et l'enseignement cherchent à faire progresser les élèves. Ce travail de terrain permet la discussion et la collaboration entre l'élève et ses pairs, et entre l'élève et l'enseignant·e.

La lecture du tableau 10.1 permet de constater que l'erreur, en ce qui a trait à la syntaxe[83], a un statut différent au sein de chacune des disciplines : outil de mesure pour des habiletés métalinguistiques en psycholinguistique, l'erreur constitue un point de référence sur le niveau de développement de l'élève pour l'enseignant. L'erreur est constamment mise à contribution dans les devis expérimentaux en psycholinguistique, notamment dans les tâches de jugement, de correction, de répétition et de réplication, qui sont les plus fréquemment utilisées dans la mesure des habiletés métalinguistiques (Gaux & Gombert, 1999a, b ; Deacon & Kieffer, 2018). À part en ce qui a trait aux activités de correction, notamment dans les exercices de cacographies (Besse, H., 2019)[84], la didactique du français n'encourage pas le travail de notions grammaticales du point de vue de l'erreur ou des pseudomots. L'un des objectifs de l'enseignement grammatical étant la maitrise de l'orthographe, on conçoit bien que les enseignant·es et les didacticien·nes n'aient pas le réflexe de recourir à des exercices de reproduction de formes grammaticales erronées pour aborder la langue, probablement de peur que les erreurs ne se fossilisent.

---

[83] On pourrait faire l'hypothèse que les épreuves psycholinguistiques ayant trait aux pseudomots, dans le cas de la conscience morphologique, sont l'équivalent morphologique de l'erreur au sein de la conscience syntaxique, en ce sens que leur inclusion dans des activités grammaticales est susceptible de contribuer à augmenter le niveau de contrôle requis pour répondre à la situation problème.

[84] Les exercices de cacographie consistent à présenter aux élèves des textes, phrases ou mots comportant des fautes sur les plans orthographique ou syntaxique, avec la consigne de les corriger. Ils ont été utilisés pour l'enseignement de la grammaire, en particulier au XIX^{e} siècle, avant d'être abandonnés au profit d'exercices risquant moins de laisser des empreintes orthographiques fautives chez les élèves.

10.4 L'erreur comme levier didactique

La tâche de réplication de l'erreur, utilisée pour mesurer les habiletés métasyntaxiques dans les devis expérimentaux en acquisition (L1 et L2) et en psycholinguistique, semble tout à fait absente des pratiques et des recommandations didactiques, même si des recherches menant à leur utilisation dans des contextes d'exercisation commencent à être mises en place (Foucambert et Redmond, 2018). Pourtant, cette tâche, transposée en activité, pourrait servir de levier à l'apprentissage grammatical, en présentant des fondements didactiques forts : la manipulation par l'élève de la grammaire (Bulea Bronckart, 2022), l'utilisation (ou non) du métalangage et le recours à des stratégies métacognitives (Chiss et David, 2018; Fisher et Nadeau, 2014; Nadeau et Fisher, 2006). Dans une perspective bien vygotskienne, cette activité grammaticale, balisée et modélisée, présente aussi l'avantage de pouvoir être réalisée en équipe. Par les interactions en classe entre élèves, et entre ceux et celles-ci et l'enseignant·e, l'activité en question permet de prendre en considération les raisonnements grammaticaux réels des élèves : elle peut ainsi faire sens pour les élèves (Roegiers, 2007). L'enseignant·e peut en effet amener les élèves à écrire des traces de leur démarche, un peu comme en mathématiques, où ce dispositif est généralisé (Kaput *et al.*, 2008).

Il nous est d'avis que la didactique vise à rendre conscient un certain travail *méta*, et à rendre disponibles les notions de contrôle, d'explicitation, d'automatisation et de raisonnement grammatical développé par modelage chez les élèves. La métalangue

a certes un rôle à jouer en classe, mais il faut que celle-ci reste une aide au raisonnement grammatical. La résolution de problèmes grammaticaux mis en scène par une activité analogue à l'épreuve de réplication de l'erreur appert représenter une occasion privilégiée de mettre véritablement les éléments grammaticaux à distance, en mettant le métalangage au service du raisonnement grammatical. L'utilisation de l'erreur en didactique de la grammaire pourrait permettre de développer chez les élèves les habiletés de reconnaissance, de manipulation et de contrôle intentionnel des structures syntaxiques, habiletés elles-mêmes constitutives des habiletés métasyntaxiques, qu'on sait prédictrices d'habiletés de haut niveau.

10.5 De l'erreur au méta, au raisonnement

Les activités sur l'erreur ou sur les pseudomots sont susceptibles de générer un étonnement, une surprise chez l'élève (Thievenaz, 2015). Nous suggérons de prendre appui sur cet étonnement pour l'envisager comme le déclencheur d'un conflit, un problème à résoudre. En effet, un exercice de grammaire basé sur la réplication d'erreurs peut être envisagé comme une situation problème, un analogue au problème arithmétique. Un problème correspond à une situation pour laquelle un individu ne dispose pas de procédure à priori pour y répondre (Chi et Glaser, 1985). La situation problème est un dispositif didactique dans lequel l'élève est mis devant une contradiction, un problème à résoudre (Fobre, 1997). La résolution du problème réside dans la propension de l'élève à reconfigurer les éléments du problème pour atteindre

une forme de stabilisation. L'exemple le plus évident est sans doute le problème arithmétique avec mise en situation dans les cours de mathématiques.

Vygotski et Piaget partagent cette idée que c'est dans une forme de contradiction qu'a lieu l'apprentissage (Baudrit, 2007). Le conflit sociocognitif, qui joue un rôle important dans les théories psychosociales et sociocognitives de l'apprentissage et du développement, fait référence à une situation dans laquelle il y a une contradiction, un entrechoquement, entre ce que l'enfant croit et ce que la réalité lui renvoie (Baudrit, 2007). Il représente un concept essentiel dans le courant socioconstructiviste néopiagétien développé par Lev Vygotski et Jerome Bruner, qui soutiennent que l'apprentissage se produit grâce aux découvertes, expériences et interactions sociales que vit l'enfant (Baudrit, 2007). Ce type de conflit émerge généralement dès l'âge de 6-7 ans, lors de la phase d'« opérations concrètes », dans la conception de Piaget (Houdé, 2023). Il s'agit du stade lors duquel les enfants sont capables de réaliser des opérations mentales de plus en plus complexes, ce qui leur permet de faire face à des problèmes dans le contexte des interactions sociales (Baudrit, 2007). Lorsque l'enfant prend conscience de cette contradiction, cela peut entrainer des déséquilibres ou des perturbations, le poussant à remettre en question ses propres connaissances, les restructurer, et à en acquérir de nouvelles (Baudrit, 2007; Bruner, 2000).

La « pause », que nous avons théorisée dans le chapitre 1, sur les habiletés métacognitives, est en quelque sorte garantie par un travail sur l'erreur et devient

presque de facto métacognitive dans une situation problème[85]. En effet, elle est autant la réponse au conflit que le signe d'une prise de distance par rapport à l'objet cognitif sur lequel diriger son attention, caractéristique de l'expérience métacognitive.

La situation problème nous semble attrayante par sa propension à demander une forme de justification grammaticale, analogue à la notion de raisonnement grammatical, essentielle en didactique de la grammaire, et pourtant peu définie. « Le manque d'études scientifiques au sujet du raisonnement des élèves dans le domaine de la langue française et de sa grammaire met en évidence un besoin d'exploration empirique et de modélisation de ce champ épistémique » (Stoudmann, communication personnelle, 2022). En effet, il semble que la notion soit mobilisée sans être explicitement définie : Fisher et Nadeau (2014, p.172) abordent la notion de « raisonnement grammatical complet »; Gauvin *et al.* (2013) reconnaissent le lien entre raisonnement grammatical et activité métalinguistique, et certains travaux invoquent le concept de « connaissances grammaticales » (Boivin, 2009, 2014; Gauvin et Thibeault, 2016; Nadeau et Fisher, 2006). Est-ce que nous pouvons envisager le raisonnement grammatical comme un type spécifique d'activité métalinguistique?

[85] La tâche de décision lexicale permet elle aussi de vérifier l'impact de la présence d'un conflit de raisonnements, grâce aux temps de réponse. En effet, les mots ayant trait aux mots cibles (reliés aux stéréotypes ayant été suscités) vont être inhibés et donc reconnus moins rapidement. Cette différence dans le temps de réponse montre que les sujets sont sensibles au conflit auquel elles et ils sont exposé·es, et tentent d'inhiber la réponse intuitive.

Il nous semble prometteur, dans un premier temps, d'envisager le raisonnement grammatical en le rapprochant de son concept d'appartenance, le raisonnement logique, de le modéliser comme tel, et de l'entrevoir comme une forme de justification dans le cadre d'une résolution de problème proche de la tâche de réplication de l'erreur. « Le raisonnement est l'activité cognitive qui consiste à évaluer des arguments. Tous les arguments impliquent l'affirmation qu'une ou plusieurs propositions (les prémisses) fournissent des raisons d'accepter une autre proposition (la conclusion) » (Goel *et al.*, 2000). La psychologie cognitive jette un regard nouveau – et différent, par rapport à celui des philosophes – sur la logique, le raisonnement et les inférences, et s'attèle à décrire les mécanismes permettant le raisonnement logique, puisque celui-ci est une partie intégrante de l'esprit humain. On sait que le raisonnement est tributaire du fonctionnement de la mémoire de travail, de l'attention et d'autres fonctions exécutives (Follmer, 2017; Johnson-Laird, 2010). De nombreux spécialistes des sciences cognitives ont supposé que le raisonnement humain dépendait de règles formelles inconscientes d'inférence. Piaget a montré que le raisonnement hypothético-déductif s'élabore progressivement chez l'enfant, à partir de 11 ans (Houdé, 2023). L'esprit humain fonctionne effectivement probablement en partie grâce à un système de déduction, et la logique traditionnelle, même si l'idée est contestée, notamment par Johnson-Laird *et al.* (2015), en partie parce que la déduction peut mener à des conclusions farfelues, et traite d'arguments conditionnels sans tenir compte des probabilités (Johnson-Laird *et al.*, 2015).

Cette idée est omniprésente dans la littérature sur les habiletés métacognitives : pour résoudre un problème, nous avons recours à des formes de raisonnement, qui peuvent, à l'occasion, s'avérer biaisées. La majorité des individus basent leur réponse à une tâche classique de raisonnement ou de prise de décision sur leurs émotions ou un autre processus plus intuitif (Evans et Stanovich, 2013; Mercier et Sperber, 2011). Les exemples concernant ces « erreurs de jugement » ou ces biais heuristiques abondent dans la littérature[86].

Ces considérations jettent les bases de l'établissement d'un lien entre la métacognition et le modèle des processus duaux à l'œuvre dans le raisonnement humain (Bigras, 2019). Les concepts de processus implicites et explicites sont définitoires du modèle des processus duaux (ou DPT, pour *dual-process theory*), théorie qui suppose l'existence de deux types de processus cognitifs mobilisés dans les tâches en psychologie cognitive (Evans et Stanovich, 2013; Kahneman et

[86] Les erreurs de raisonnement, comme on en trouve dans *la Epstein's Ratio Bias Task*, arrivent quand il y a un conflit entre un raisonnement logique et un raisonnement de type intuitif. Un exemple emblématique d'un raisonnement fautif (probabilistique, ici) est la tâche de Epstein (1994) (*Epstein's Ratio Bias Task*). Dans cette tâche, deux plateaux de *jelly beans* sont présentés aux participant·es : le premier contient un *jelly bean* rouge et neuf *jelly beans* blancs, et le deuxième contient neuf *jelly beans* rouges, sur un total de 100. Il s'agit de déterminer le plateau à sélectionner pour maximiser les chances de piger au hasard un *jelly bean* rouge. D'un point de vue logique et probabiliste, il est évident que la sélection du premier plateau est la conclusion correcte dans cette tâche de prise de décision. Pourtant, la majorité des gens vont échouer à établir le bon ratio et opter pour le tableau où les jelly beans rouges sont les plus nombreux, succombant à un raisonnement intuitif et rapide. Ce type d'expérience tend à conforter les tenants d'une position forte sur les processus duaux dans (Evans, & Stanovich, 2013). Certaines erreurs de raisonnement, notamment les biais de croyance, peuvent être en partie expliquées par le recours au système 1.

Frederick, 2002; Thompson *et al.*, 2011); à savoir un premier, réputé simple, automatique, implicite et efficace, relevant de biais et d'heuristiques (Karimi et Ferreira, 2016) (système 1), et le deuxième, constituant un système de pensée analytique, délibérée, explicite et de détection d'erreurs (système 2) (Bigras, 2019).

Des tâches telles que la réplication de l'erreur présentent un conflit, une complication pour lesquels le raisonnement heuristique (système 1) n'est pas pertinent. Celles-ci forcent le recours à un type de raisonnement plus analytique, plus délibéré : plus *méta*, en somme. Ces tâches mesurent finalement la capacité des sujets à, d'une part, détecter le conflit, c'est-à-dire une situation dans laquelle un raisonnement intuitif sera en contradiction avec un raisonnement probabiliste ou logique, et d'autre part, à vérifier leur comportement à la suite de la détection du conflit. Le raisonnement rapide, heuristique, et éventuellement fautif, est inhibé, et peut donner lieu à une pensée qui permet de dresser une réponse adaptée à une situation nouvelle.

10.6 Dispositif didactique proposé

La présente réflexion sur la possible transposition de tâches métalinguistiques mettant en œuvre l'erreur en activités scolaires et l'importance des représentations linguistiques dans la réussite aux épreuves métasyntaxiques, constat que nous avons établi dans nos analyses et notre discussion des résultats, aux chapitres VIII et IX, nous amènent vers la question de la mise en place des savoirs à enseigner en classe de français.

> « Plus que dans d'autres disciplines internes à la matière « français », l'habituation des formés à l'activité réflexive, métalinguistique suppose de problématiser des données de langue au lieu d'en fournir immédiatement la description réglée, de s'interroger sur l'appareil notionnel existant et donc de prévoir, à côté des leçons magistrales, des démarches effectives de recherche apprenant à poser un problème à partir d'un corpus, d'une règle, d'une définition et à réunir les conditions permettant de le résoudre. » (Chiss et David, 2001, p.101)

L'enseignement métasyntaxique, qui implique différents niveaux de collaboration, d'explicitation, de complexité, différents objectifs, doit pouvoir avoir un impact sur le développement du raisonnement grammatical chez les élèves. Nous faisons la proposition suivante : la réplication d'erreurs constitue un moyen de pointer l'objet grammatical chez les élèves, sans nécessairement le nommer. Il permet ainsi de contourner, pour un temps, l'utilisation du métalangage, parfois rédhibitoire (Bronckart et Chiss, 1990), en mettant en place un dispositif[87] où l'élève doit se focaliser sur le rôle que joue l'élément linguistique en question dans la phrase, dans le texte.

L'enseignant·e, en tant qu'animateur, animatrice de l'atelier, lance par la suite un débat métalinguistique, dans lequel les élèves sont amené·es à proposer leurs raisonnements, en adoptant de facto une posture métalinguistique, en explicitant des éléments implicites, qui appartiennent à l'épilinguistique (Chiss et David, 2011).

---

[87] Un dispositif didactique est entendu ici comme un ensemble de moyens permettant « l'activation d' **"**outils de médiation**"** (albums et textes, fiches, écrits comme aide-mémoire, constats, voix de l'enseignant) servant à mettre en relation les élèves avec l' **"**objet**"**, à guider l'attention sur cet objet » (Blanc *et al.*, 2023).

L'activité proposée est conçue pour être travaillée en petits groupes d'élèves, dans le cadre de modules grammaticaux ciblés, une disposition que permet la situation problème (Astolfi *et al.*, 2008). L'intérêt d'une telle activité est de placer les élèves dans une posture de chercheur·e (Astolfi *et al.*, 2008). Dans l'épreuve de déplacement, l'une des tâches de CS utilisées dans le cadre de cette thèse et décrites dans la section 6.2.2, l'élève doit répliquer une erreur par analogie : il ou elle doit en effet avoir préalablement compris l'erreur systémique ayant eu lieu dans la phrase précédente, et appliquer cette compréhension du système dictant l'erreur dans la phrase suivante. Partant, on peut imaginer une activité grammaticale de l'ordre de celle figurant dans le tableau 10.2.

Tableau 10.2 Proposition d'activité grammaticale sur le complément du nom mettant en jeu la réplication de l'erreur

Analyse du complément du nom

Âge des élèves : 9-15 ans

Objectifs :

- Constater que le CdN est un complément facultatif, mais sémantiquement riche, du GN qui dépend syntaxiquement du noyau, le nom.
- Constater le caractère hiérarchique des constituants dans la phase.

Exercice :

Yoda, comme tu le sais, parle une langue bien à lui. Peux-tu reproduire sa façon de parler (Y) dans les phrases cible A et B?

Y : Mon frère suit <u>de natation</u> des cours.

Phrase cible A : *Son ami veut une tasse <u>de café</u>.*

Phrase cible B : *Le jeune garçon porte une veste <u>de laine angora</u>.*

Phase d'institutionnalisation :

- Quel est le raisonnement qui permet de trouver les réponses ?
- Les phrases A et B ont la même structure. Peux-tu lister les constituants/faire un arbre syntaxique de ces phrases ?
- Sur la base de la réponse qui précède, peux-tu lister les constituants/faire un arbre syntaxique de la façon de parler de Yoda ?
- À quoi peut servir le CdN ?

Cet exemple d'activité est construit en partie sur la base d'une proposition linguistique et méthodologique forte et éprouvée en didactique de la grammaire : les manipulations syntaxiques (Bulea Bronckart, 2022; Chartrand, 2012a; Fisher et Nadeau, 2014). Les manipulations syntaxiques sont des tests opératoires qui n'ont pas été conçus comme des objets de savoir (Chartrand, 2013; Bulea Bronckart, 2022). Ce sont des traitements transformationnels qui sont une spécificité de la grammaire. Ils ont pour objectif de contribuer à construire, « chez les élèves, […] une représentation systémique de la langue et au développement des capacités de mise en pratique de la langue dans des textes » (Bulea Bronckart, 2022, p. 1). La transposition de savoirs, de concepts, doit être accompagnée d'une transmission de méthodes, d'opérations et de savoir-faire (Chartrand, 2012a). Nous prônons un usage développé et renforcé de ce type de manipulations syntaxiques en proposant des activités de transformations d'énoncés « sans erreurs », ou agrammaticaux, en énoncés « fautifs », ou en

énoncés comportant des erreurs, sans que l'outil ne devienne objet (Arseneau *et al.*, 2016). Ceci, bien sûr, dans le but de faire apparaitre des notions grammaticales sur lesquelles les élèves s'attèlent à réfléchir (Bulea Bronckart, 2022), sur la base de leurs connaissances épilinguistiques d'abord, puis en mettant en place un raisonnement grammatical, ou un raisonnement donnant une large place au métalinguistique (Audion, 2016).

10.7 Synthèse des perspectives didactiques en lien avec un dispositif basé sur l'erreur

Ce chapitre à perspectives didactiques avait pour ambition d'aborder le même construit à priori, soit les habiletés métasyntaxiques, et de l'aborder sous deux angles : celui de la psycholinguistique et celui de la didactique. Nous avons effectué un examen comparé des différences épistémologiques et méthodologiques entre les deux disciplines, par un travail de réflexion sur les concepts sous-jacents à leurs outils d'évaluation et d'exercisation, avec l'objectif de répondre à la question suivante :

- *Comment peut-on opérer le passage entre tâches métalinguistiques et activités grammaticales?*

Nous avons d'abord passé en revue les activités grammaticales, telles qu'elles sont proposées en didactique du français. Elles ont été analysées à travers le prisme des dimensions issues de la reconceptualisation des épreuves métalinguistiques, un

travail notamment effectué par Bialystok (2001a) et présenté au chapitre I. Nous avons ensuite examiné le traitement potentiel des épreuves recensées dans la littérature psycholinguistique de manière à penser leur transposition en activités didactiques, explicites par nature (Zorman *et al.*, 2015). Cette tentative de rapprochement entre tâches psycholinguistiques et activités scolaires nous a permis de relever les divergences épistémologiques dans les deux disciplines, mais aussi des objectifs communs.

Ces prémices nous ont permis d'énoncer et de clarifier la nature épistémologique du construit *méta* dans les deux domaines, ce qui constituait le deuxième objectif de ce chapitre. Finalement, nous avons abordé quelques pistes de réflexion sur l'enseignement grammatical, au regard des habiletés métalinguistiques, qu'il est censé mobiliser. Nous avons émis quelques propositions pour une transposition didactique heureuse des tâches psycholinguistiques aux activités grammaticales, notamment par l'aménagement d'activités novatrices en didactique du français qui donnent une large place à l'erreur syntaxique et morphologique, l'idée étant de tester des dispositifs didactiques conçus sous la forme de situations problèmes pour favoriser l'émergence puis le développement du raisonnement grammatical. Nous avons proposé que les tâches de réplication de l'erreur puissent être utilisées en classe comme un moyen d'automatiser les capacités qui peuvent conduire à de meilleures habiletés métasyntaxiques, et ultimement, à une meilleure compréhension

des textes. En passant par une phase de résolution de problème, les élèves sont encouragés à réfléchir sur leur langue et à observer son fonctionnement.

Quatre éléments nous semblent importants à retenir. En premier lieu, nous sommes d'avis que la démonstration des différences entre les opérations cognitives et les niveaux d'explicitation des connaissances mobilisées dans l'ensemble des activités grammaticales atteste de l'importance de l'explicitation et de la justification grammaticale afin de développer le raisonnement grammatical chez les élèves. Un enseignement explicite de justifications d'ordre grammatical permet de rendre conscients des concepts, des opérations et des savoir-faire pour qu'ils deviennent des objets de savoir pour l'élève.

En deuxième lieu, l'examen approfondi des activités grammaticales et des habiletés sous-jacentes qu'elles mettent en œuvre devrait permettre une catégorisation de celles-ci, ou du moins, une réflexion certaine, par les enseignant·e·s et les formatrices et formateurs. Cette nécessaire prise de conscience de la part des enseignant·e·s pourrait permettre de mieux cibler les difficultés des élèves afin d'améliorer l'accompagnement dont ils et elles bénéficient dans leur cheminement grammatical. Cet effort de reconceptualisation permet donc de repenser les gestes enseignants et le pointage d'informations pertinentes permettant de déclencher des processus de décomposition et de recomposition.

En troisième lieu, nous souhaitons retenir l'idée de reprendre des tâches psycholinguistiques mobilisant des habiletés purement linguistiques (qui correspondent aux épreuves appartenant au quadrant en bas à gauche dans le modèle de Bialystok (2001a), représenté à la figure 1.1) et de les travailler dans le cadre de la classe. Des exercices répétitifs comme le test de closure ou la détection d'erreurs pourraient favoriser l'automatisation d'habiletés linguistiques qui président au développement d'habiletés linguistiques plus complexes.

En quatrième et dernier lieu, nous plaidons pour un rapport plus décomplexé des enseignant·es et des chercheur·es vis-à-vis de l'erreur, telle que comprise au sein de cette thèse, et surtout pour son inclusion dans un dispositif didactique visant à travailler le raisonnement grammatical. Plus spécifiquement, nous sommes d'avis que l'enseignement de la grammaire, lorsqu'il n'est pas au service de l'enseignement de l'orthographe, mais sert surtout une conception abstraite de la langue, peut passer par des phases d'exercisation lors desquelles la consigne de base serait de chercher à reproduire des erreurs détectées sur une phrase qui est grammaticale (Foucambert et Redmond, 2018). Nous réitérons que pour Gaux et Gombert (1999b), c'est l'épreuve de réplication de l'erreur qui permet d'évaluer les habiletés métasyntaxiques, puisque seuls une réflexion et un contrôle intentionnel sur la langue permettent de la réussir. Il est concevable que l'idée d'une reproduction de l'erreur choque, puisque les enseignant·es sont certainement attaché·es à l'idée d'éviter les erreurs. L'exercice touche en effet à une conception du fonctionnement de la langue

qui est dans l'abstraction. Nous sommes d'avis que cette épreuve et l'activité transposée et didactisée qui lui serait analogue permettent une véritable compréhension des phénomènes grammaticaux, qui se situe au-delà de la correction, et juste avant l'explicitation de règles grammaticales formelles. Il faut bien concevoir qu'une telle activité, par ce travail de reconceptualisation de l'erreur, sera au service de la réduction des erreurs.

À travers cette démonstration des ponts et des contrastes entre psycholinguistique et didactique, il est possible de réfléchir à de meilleures combinaisons de tâches. L'étonnement, accompagné de manipulations syntaxiques menant à un raisonnement grammatical (Bulea Bronckart, 2022), pourrait être considéré comme déclencheur de processus métacognitifs dans l'enseignement de la grammaire, comme nous espérons l'avoir démontré, mais aussi, peut-être, dans la formation des enseignant·es (Thievenaz, 2015). Nous proposons de décloisonner la mesure et l'exercisation des habiletés métasyntaxiques, en intégrant éventuellement l'idée de tenir compte de l'utilisation textuelle de la grammaire en psycholinguistique, et retenons l'idée de considérer la reproduction de l'erreur comme agent d'apprentissages grammaticaux en didactique de la grammaire.

CONCLUSION

Les résultats étayés dans le cadre de ce projet éclairent une portion supplémentaire de la recherche sur les processus de lecture et sur les habiletés linguistiques et cognitives susceptibles d'engendrer une lecture experte. Nous avions comme ambition, par une véritable tentative de déconstruction théorique et statistique des tâches utilisées dans la littérature sur les habiletés métalinguistiques, d'ouvrir une discussion sur les fonctions cognitives que les épreuves utilisées pour mesurer les différentes habiletés métalinguistiques viennent réellement mobiliser. Les retombées sont de trois ordres : théorique, méthodologique (docimologique) et didactique.

Dans une perspective épistémologique, nous avons apporté notre pierre à l'édifice du grand projet de clarification du construit « méta », dont les acceptions au sein de la psycholinguistique et de la didactique sont variées et poreuses, dans l'espoir de créer des ponts entre les disciplines et de favoriser les échanges conceptuels et scientifiques entre elles.

En remettant en question l'unicité des tâches traditionnellement utilisées pour évaluer les habiletés métalinguistiques, cette étude a contribué à mettre en évidence la nécessité de repenser les approches méthodologiques et docimologiques utilisées. Notre recherche souligne l'importance de prendre en compte les unités linguistiques

et les processus cognitifs sous-jacents à ces habiletés lors de la conception des épreuves d'évaluation. En adoptant une approche plus rigoureuse et holistique des mesures, nous pourrons mieux comprendre les mécanismes impliqués dans la lecture experte et mettre en place des outils d'évaluation plus efficaces. L'un des apports méthodologiques de la thèse est également cette critique du caractère imbriqué des épreuves de CS, que nous n'avons pu prendre en compte dans notre analyse. Ces implications méthodologiques offrent ainsi des perspectives prometteuses pour le développement de futurs travaux de recherche.

Sur le plan didactique, nous avons postulé que les résultats de notre recherche, de nature psycholinguistique, nous permettaient d'envisager une opération de retournement didactique par transposition de la tâche en contexte d'expérience à l'activité en contexte de classe. Nous avons procédé à une reconceptualisation de construits présentant un intérêt pour les didacticien·nes et le monde de l'Éducation, en essayant de comprendre plus finement les habiletés métalinguistiques qui influencent la compréhension en lecture chez une population préadolescente. Nous nous positionnons ainsi pour des sciences de l'éducation qui permettent de rendre les élèves autonomes dans leur apprentissage. Les habiletés de résolution de problèmes peuvent être travaillées et objectivées, afin de rendre les élèves actives et actifs dans leur apprentissage.

De plus, d'un point de vue expérimental, et dans une perspective de réhabilitation de données issues de la recherche psycholinguistique et didactique vers la formation des enseignant·es et le « terrain » (Bressoux, 2017; Bronckart et Chiss, 1990), nos résultats ont permis d'établir plus clairement les liens qui se nouent entre les habiletés métasyntaxiques et la compréhension en lecture. Alors que les manuels scolaires et les plans d'étude distinguent grammaire et compétences langagières, alors que l'une des finalités de l'enseignement grammatical concerne le développement d'habiletés de haut niveau, il est heureux de constater que ces deux jeux d'habiletés se construisent et se nourrissent l'un l'autre. Ils participent tous deux d'une accession pertinente, systémique aux représentations de la langue et son fonctionnement, et d'une compréhension du monde également. Comme l'écrivent Chiss et David (2011, p. 101), « explorer le fonctionnement d'une langue ce n'est donc pas uniquement se l'approprier, c'est également prendre conscience que cette exploration transforme l'individu dans son rapport au monde et à l'autre ».

Les perspectives de recherche futures sont nombreuses et foisonnantes dans les champs des habiletés métalinguistiques et de la compréhension en lecture, tant ces construits sont perméables à l'influence de plusieurs épistémologies et les nourrissent. Nous avons déjà fait état des biais méthodologiques contenus dans la présente thèse et des moyens d'y répondre afin de mener à bien des analyses complémentaires. Tout d'abord, il serait intéressant d'approfondir nos connaissances sur les aspects spécifiques qui n'ont pas été pleinement explorés dans cette étude.

Un regard plus pointu sur les réponses, notamment erronées, des élèves à des items en particulier, par exemple à des fonctions syntaxiques particulières, pourrait également permettre de lever le voile sur les raisonnements grammaticaux des élèves et sur les éléments linguistiques qui leur posent le plus problème. De plus, il serait judicieux de poursuivre les recherches en utilisant des méthodes d'analyse différentes, telles que des études longitudinales ou en incluant des variables contrôle, afin de renforcer la validité des résultats. D'autres recherches seront nécessaires afin de clarifier l'impact des dimensions sous-jacentes dans les épreuves de conscience métalinguistique sur les habiletés langagières de haut niveau et d'évaluer les processus et connaissances qu'elles mettent en œuvre. Enfin, il serait bénéfique d'élargir la portée de cette recherche en incluant des populations variées, notamment des sujets en difficultés d'apprentissage ou des élèves issus de milieux socioéconomiques divers. Les possibilités de recherche futures sont vastes, passionnantes, et nous faisons le vœu que cette thèse serve de point de départ pour de nouvelles investigations.

ANNEXE A

PROTOCOLE DE PASSATION, ÉPREUVES DE CONSCIENCE MORPHOLOGIQUE

Matériel expérimental des différentes épreuves, vue générale

- **6 épreuves sur les mots de même famille dans Logix**
- **1 épreuve de compréhension en lecture avec petits textes dans Psychopy**
- **3 épreuves de réplication de l'erreur dans Psychopy**
- **2 épreuves matricielles dans Psychopy**
- **1 épreuve de compréhension en lecture avec long texte sur papier**

Déroulement

Première séance

- **Épreuve avec paires de mots (ou de jugement de relations de mots) – Épreuve n°1**
- **Épreuve d'extraction de la base - Épreuve n°2**
- **Épreuve de dérivation - Épreuve n°3**

Deuxième séance

- **Épreuve de décomposition - Épreuve n°4**
- **Épreuve de choix d'affixes dérivationnels dans des pseudomots - Épreuve n°5**
- **Épreuve de choix d'affixes flexionnels dans des pseudomots - Épreuve n°6**

Troisième séance

- **Matrices**
- **Classification**
- **Épreuve de compréhension en lecture sur courts textes**

Quatrième séance

- **Remplacement**
- **Addition suppression**
- **Déplacement**

Cinquième séance

- **Épreuve de compréhension en lecture sur texte long**

La présence en classe de l'expérimentatrice ou de l'expérimentateur, et de l'enseignant ou l'enseignante, s'il y a lieu, est requise pour assurer le bon fonctionnement de l'expérience. Ils ne doivent pas répondre aux questions des élèves pendant la passation, si celles-ci ont trait aux items. Ils peuvent toutefois répondre aux questions des élèves si un problème technique se présente (bogues des logiciels, etc.). Les élèves sont autorisés à poser des questions et à obtenir des réponses portant sur le fonctionnement de l'épreuve ou sur des items en particulier seulement lors de la période d'apprentissage ou de pratique. Toutes les épreuves sont précédées d'une séance dite « d'apprentissage », dans laquelle les élèves apprennent à se familiariser avec la nouvelle épreuve présentée et avec le fonctionnement du logiciel.

Il est à noter que les deux logiciels enregistrent les temps de réponse et la qualité de la réponse, ou la réussite des participantes et participants à chaque item présenté. Il est primordial de s'assurer que les élèves ont compris l'importance de garder le silence pendant la passation des épreuves, pour assurer un environnement de travail équivalent à chacun des participants. Les élèves qui terminent une épreuve avant les autres doivent pouvoir s'occuper de façon silencieuse. Il faut s'assurer que tous les élèves ont terminé l'épreuve avant de passer à la présentation de l'épreuve suivante. Si certains élèves mettent trop de temps à finaliser l'épreuve, notamment celles de dérivation et de décomposition, il est recommandé d'écrire XX dans le logiciel, puis de taper sur Entrée. Ces items recevront un traitement différé lors de l'analyse des résultats.

Pendant l'épreuve, les élèves ont été encouragés à se faire confiance et à répondre le plus rapidement et le mieux possible.

Première séance

LOGIX©

Épreuve avec paires de mots (ou de jugement de relations de mots) – Épreuve n°1

Consignes dans le logiciel

Consignes :##Certains mots sont de la même famille.##Par exemple, école et écolier sont des mots de la même famille. Les mots curieux et curiosité sont de la famille.##Par contre, les mots bouche et boucherie se ressemblent, mais ne sont pas de la même famille. Ils n'ont pas de sens en commun.###Deux mots seront présentés ensemble à l'écran.##Appuie sur O, pour oui, quand ils sont de la même famille.##Appuie sur N, pour non, quand ils ne sont pas de la même famille.

Présentation

Sais-tu ce qu'est un mot de la même famille? Par exemple, *jardin, jardiner, jardinier, jardinière, jardinet, jardinage*, ce sont tous des mots de la même famille : ils parlent tous de jardin. Un autre exemple : *pommier*, c'est un mot de la même famille que *pomme*, parce que les pommes, ça pousse dans les pommiers. L'arbre s'appelle le pommier. Maintenant, *poisson*, est-ce que c'est un mot de la même famille que *pois* (comme un petit pois vert)? Non! *Pois* n'est pas de la même famille que poisson. C'est juste un accident si les deux mots commencent par le même son.

Dans cet exercice, je vais te présenter deux par deux des mots qui se ressemblent un peu. Certains de ces mots sont des mots de la même famille, d'autres pas. Je veux que tu me dises chaque fois si les mots que je te dis sont de la même famille ou non.

On va s'entrainer un peu. Tu dois appuyer sur la lettre O, pour oui, quand les mots sont de la même famille. Tu dois appuyer sur la lettre N, pour non, quand ils ne sont pas de la même famille.

Clique sur Apprentissage. On peut lire les consignes. (lire les consignes ci-haut) Clique sur Continuer. Le premier exemple est « **école écolier** ». Tu dois appuyer sur la lettre O parce que les deux mots sont de la même famille. Le mot écolier est un mot de la même famille que école. L'écolier va à l'école.

Le 2e exemple est « **heure heureux** ». Est-ce que tu penses que ces deux mots sont de la même famille? Pourquoi? (Interaction avec les élèves) L'heure n'a rien à voir avec le mot heureux. Ce ne sont pas des mots de la même famille. C'est juste un accident si les deux mots commencent par le même son. Tu dois donc appuyer sur N pour « NON ».

Maintenant, regarde le 3e exemple : « **dent dentiste** ». Tu appuies sur la lettre O ou sur la lettre N ? Bravo, tu as raison! Le mot dent et le mot dentiste sont de la même famille. Un dentiste, c'est une personne qui répare les dents.

Le dernier exemple est « **bague baguette** ». Est-ce que tu penses que ces deux mots sont de la même famille? Pourquoi? (Interaction avec les élèves). Ce ne sont pas des mots de la même famille. Une bague, ça n'a rien à voir avec une baguette, comme une baguette de magicien. Encore une fois, c'est juste un accident si les deux mots commencent par les mêmes lettres. Il faut que tu appuies sur N.

As-tu des questions sur ce que tu dois faire? Est-ce que tu as compris? Es-tu prêt ou prête? Clique sur le bouton « Épreuve ». Essaie de répondre vite et bien.

Épreuve d'extraction de la base - Épreuve n°2

Consignes logiciel

Consignes :##Un mot apparait en haut de l'écran.##Parfois, ce ne sera pas un vrai mot, mais ce n'est pas grave!##Essaie de trouver un mot plus petit à l'intérieur de ce mot.##Le mot que tu écris doit pouvoir être de la même famille que le mot présenté.##Si tu ne trouves pas de mot plus petit, réécris le mot présenté.##Appuie sur entrée pour passer au prochain mot.

Présentation

On a vu plus tôt des mots qui étaient de la même famille. On avait dit que par exemple, *jardin*, *jardiner*, *jardinage* ou *jardinet*, ce sont des mots de la même famille : ils parlent tous de jardins. Maintenant, ce qu'on va faire, c'est essayer de trouver un plus petit mot à l'intérieur d'un mot. Il faut que les deux soient de la même famille.

Par exemple, dans le mot fillette, il y a le mot fille. Le mot plus petit à l'intérieur du mot fillette est le mot fille. Ils sont de la même famille, parce qu'une fillette, c'est une petite fille. Le mot que tu auras à écrire est le mot fille.

Clique sur Apprentissage. On peut lire les consignes. (Lire les consignes). Maintenant, clique sur Continuer.

On va commencer l'exercice avec le mot **tassette**. *Tassette*, ce n'est pas un vrai mot! Mais s'il existait, on pourrait penser que c'est une petite tasse. Je pourrais dire que je bois mon café dans une tassette, par exemple! Le mot plus petit à l'intérieur du mot tassette est le mot tasse. *Tasse* semble aussi être dans la même famille que t*assette*. Alors tu tapes le mot tasse et tu appuies sur Entrée.

Par la suite, c'est le mot **alouette** qui apparait. Est-ce qu'il y a un mot plus petit dans le mot *alouette*? Dans le mot *alouette*, il n'y a pas de mot plus petit qui est de la même famille. Tu réécris le mot alouette et tu appuies sur Entrée.

Ensuite, c'est le mot **refaire** qui apparait. Est-ce qu'il y a un mot plus petit dans le mot *refaire*, et qui est de la même famille? Dans le mot *refaire*, il y a le mot *faire*. Le mot plus petit à l'intérieur du mot *refaire* est le mot *faire*. Ils sont de la même famille parce que refaire quelque chose, c'est faire quelque chose une deuxième fois. Alors écris le mot *faire* et appuie sur Entrée.

Après, on a le mot **renifler**. Est-ce qu'il y a un mot plus petit dans le mot *renifler*? Qu'est-ce que ça veut dire, le mot *renifler*? (Interaction avec les élèves) Est-ce que c'est comme pour le mot *refaire*, dans lequel il y a le mot faire? Dans le mot *renifler*, il

n'y a pas de mot plus petit de la même famille. Alors écris le mot *renifler* et appuie sur Entrée.

On a maintenant le mot **patinage**. Est-ce qu'il y a un mot plus petit dans le mot *patinage*? Qu'est-ce que ça veut dire, le mot *patinage*? (Interaction avec les élèves) Dans le mot *patinage*, il y a les mots *patin* et *patiner* ou *patine*. Les mots *patin* ou *patiner* sont de la même famille que le mot *patinage*. Alors tu écris le mot *patin* ou le mot *patiner* et tu appuies sur Entrée.

Le dernier mot dans notre exercice est le mot **partirage**. Ce n'est pas un vrai mot! Est-ce que tu penses qu'il y a un mot plus petit dans le mot *partirage*? (Interaction avec les élèves) Dans le mot *partirage*, on dirait qu'il y a le mot *partir*. On pourrait penser que *partirage*, c'est l'action de partir. Alors tu écris le mot *partir* et tu appuies sur Entrée.

Quand tu es prêt ou prête, tu peux appuyer sur le bouton Épreuve. Essaie de répondre vite et bien.

Épreuve de dérivation - Épreuve n°3

Consignes logiciel

##Un mot apparait en haut de l'écran.##Il y a une phrase juste en dessous, mais il manque un mot dedans!##Tu dois compléter la phrase en écrivant un mot qui ressemble à celui qui est en haut de l'écran, et qui rend la phrase correcte.##Quand tu as fini d'écrire ton mot, appuie sur Entrée.

Présentation

On va faire un autre exercice avec les mots de la même famille. Tu te rappelles, par exemple, que *jardin* et *jardiner* sont des mots de la même famille, et que *pomme* ou *pommier* sont des mots de la même famille. D'accord.

Clique sur Apprentissage. On peut lire les consignes. (Lire les consignes) Maintenant, clique sur Continuer.

Il y a un mot en haut de l'écran, et une phrase en dessous. Mais la phrase n'est pas terminée! Il faut que tu la termines en écrivant un mot, un seul mot. Ce mot doit être de la même famille que le mot qui apparait en haut de l'écran.

On va s'exercer un peu. Le premier mot est « **garage**». La phrase à compléter est « Gaston est un … ». Quel mot de la même famille que garage rendrait la phrase correcte? Oui! Le mot *garagiste*! Ça donnerait la phrase « Gaston est un garagiste ». Écris le mot *garagiste* et appuie sur Entrée.

On va faire un autre exemple. Le mot qui apparait en haut est « **délice**». La phrase à compléter est « Le repas est …». Quel mot de la même famille que *délice* rendrait la phrase correcte? Très bien! Le mot *délicieux*! Ça donnerait la phrase « Le repas est délicieux». Écris le mot *délicieux* et appuie sur Entrée.

On fait un troisième exemple. Le mot qui apparait est « **conducteur**». La phrase à compléter est « Adélie est une bonne …». Adélie est un nom de fille ou de femme. Quel mot de la même famille que *producteur* rendrait la phrase correcte? Très bien! Le mot *productrice*! Ça donnerait la phrase « Adélie est une bonne productrice». Écris le mot *productrice* et appuie sur Entrée.

On fait un dernier exemple. Le mot qui apparait est « **parler** ». La phrase à compléter commence par « Nous …». Qu'est-ce qui manque ici? Quel mot rendrait la phrase

correcte? Très bien! Le mot *parlons*! Ça donnerait la phrase « Nous parlons ». Écris le mot *parlons* et appuie sur Entrée.

Quand tu es prêt ou prête, tu peux appuyer sur le bouton Épreuve. Essaie de répondre vite et bien.

Deuxième séance

Épreuve de décomposition - Épreuve n°4

Consignes logiciel

##Un mot apparait en haut de l'écran.##Il y a une phrase juste en dessous, mais il manque un mot dedans!##Tu dois compléter la phrase en écrivant un mot qui ressemble à celui qui est en haut de l'écran, et qui rend la phrase correcte.##Quand tu as fini d'écrire ton mot, appuie sur Entrée.

Présentation

On va faire un autre exercice avec les mots de la même famille. Tu te rappelles par exemple que *jardin* et *jardiner* sont des mots de la même famille. D'accord.

Clique sur Apprentissage. On peut lire les consignes. (Lire les consignes) Maintenant, clique sur Continuer.

Il y a un mot en haut de l'écran, et une phrase en dessous. Mais la phrase n'est pas terminée! Il faut que tu la termines en écrivant un mot, un seul mot. Ce mot doit être de la même famille que le mot qui apparait en haut de l'écran.

On va s'exercer un peu. Le premier mot est « **chanteur** ». La phrase à compléter est « Mon grand frère adore … ». Quel mot de la même famille que *chanteur* rendrait la phrase correcte? Oui! Le mot *chanter*! Ça donnerait la phrase « Mon grand frère adore chanter ». Écris le mot *chanter* et appuie sur Entrée.

On va faire un autre exemple. Le mot qui apparait est « **douceur**». La phrase à compléter est « Le poil du chien est très …». Quel mot de la même famille que *douceur* rendrait la phrase correcte? Très bien! Le mot *doux*! Ça donnerait la phrase « Le poil du chien est très doux ». Écris le mot *doux* et appuie sur Entrée.

On fait un troisième exemple. Le mot qui apparait est « **animatrice**». La phrase à compléter est « Gaston est un bon …». Gaston, c'est un nom de monsieur! Quel mot de la même famille que *animatrice* rendrait la phrase correcte? Très bien! Le mot *animateur*! Ça donnerait la phrase « Gaston est un bon animateur». Écris le mot *animateur* et appuie sur Entrée.

On fait un dernier exemple. Le mot qui apparait est « **chassons** ». La phrase à compléter est « Il faut … ». Qu'est-ce qui manque ici? Quel mot de la même famille que *chassons* rendrait la phrase correcte? Très bien! Le mot *chasser*! Ça donnerait la phrase « Il faut chasser ». Écris le mot *chasser* et appuie sur Entrée.

Quand tu es prêt ou prête, tu peux appuyer sur le bouton Épreuve. Essaie de répondre vite et bien.

Épreuve de choix d'affixes dérivationnels dans des pseudomots - Épreuve n°5

Consignes logiciel

Consignes : ##Une phrase va apparaitre.#Dans cette phrase, il y a un mot qui manque.##Une fois que tu as lu cette phrase, clique sur Continuer.##On te demande alors de choisir la meilleure réponse entre les trois mots présentés.#Ce ne sont pas de vrais mots, mais ce n'est pas grave!##Tu dois cliquer sur le mot qui rend la phrase correcte.

Présentation

Dans ce nouvel exercice, tu vas voir apparaitre plusieurs phrases. Dans chacune des phrases, il va y a avoir un mot qui manque. Pour trouver le mot qui manque, tu vas devoir choisir entre 3 mots. Mais attention! Ce ne sont pas de vrais mots! Parmi les trois mots, un seul rend la phrase bien construite. Les deux autres « sonnent bizarre ». Il faut que tu choisisses le mot qui sonne le mieux, qui rend la phrase correcte.

Clique sur Apprentissage. On peut lire les consignes. (Lire les consignes) Maintenant, clique sur Continuer.

On va faire un exemple ensemble.
Regarde bien :
J'aimerais que mon chaton arrête de _____.
Clique sur Continuer. Tu as trois choix.
bluvable
bluvage
bluvir

Même si ce sont des mots inventés, tu vas voir qu'il y a en un qui fonctionne mieux avec le reste de la phrase, qui la rend correcte. Tu remarques que les trois choix commencent par les mêmes lettres. B-L-U-V.

Est-ce que tu remarques que les deux premiers choix sont bizarres alors que le troisième choix, *J'aimerais que mon chaton arrête de bluvir*, ressemble presque à une phrase normale, que tu pourrais entendre tous les jours ou dire toi-même? On pourrait dire « J'aimerais que mon chaton arrête de grossir », par exemple. *Grossir*, ça ressemble à *bluvir*! C'est le bon choix. Clique sur le mot *bluvir*, puis sur Continuer.

Maintenant, la phrase est « Mon petit frère est de plus en plus____. » Clique sur Continuer pour voir les choix. Il faut choisir la meilleure réponse. Tu remarques que les trois choix commencent par les mêmes lettres. P-R-I-L-U-S.

(Interaction avec les élèves) Qu'est-ce que tu en penses? Si tu penses que c'est le 2e mot qui est le bon, celui du milieu, tu as raison. *Prilusant*, c'est le mot qui rend la phrase bien construite. On pourrait dire « Mon petit frère est de plus en plus fatiguant », par exemple.

Mon petit frère est de plus en plus
Prilusage
Prilusant
Prilusir

Clique sur Continuer. La prochaine phrase est « Hubert et son papa vont à la ___ ». Clique sur Continuer pour voir les choix.

louviterie
louvitir
louvitable

(Interaction avec les élèves) As-tu trouvé le bon mot? Si tu as choisi le premier mot entre les trois 3 choix, le mot *louviterie*, tu as raison! C'est la fin du mot qui te fait choisir *louviterie*, puisque le début est le même. Clique sur *louviterie*, puis sur Continuer.

Tu es maintenant prêt ou prête pour l'Épreuve! Clique sur Épreuve. Essaie de répondre vite et bien.

Épreuve de choix d'affixes flexionnels dans des pseudomots - Épreuve n°6

Consignes logiciel

Consignes : ##Une phrase va apparaitre.#Dans cette phrase, il y a un mot qui manque.##Une fois que tu as lu cette phrase, clique sur Continuer.##On te demande alors de choisir la meilleure réponse entre les quatre mots présentés.#Ce ne sont pas de vrais mots, mais ce n'est pas grave!##Tu dois cliquer sur le mot qui rend la phrase correcte.

Présentation

Dans ce nouvel exercice, tu vas voir apparaitre plusieurs phrases. Dans chacune des phrases, il va y avoir un mot qui manque, comme dans l'exercice précédent. Pour trouver le mot qui manque, tu vas devoir choisir entre 4 mots. Mais attention! Ce ne sont pas de vrais mots! Parmi les trois mots, un seul rend la phrase bien construite. Les deux autres sonnent bizarre. Il faut que tu choisisses le mot qui sonne le mieux, qui rend la phrase correcte.

Clique sur Apprentissage. On peut lire les consignes. (Lire les consignes) Maintenant, clique sur Continuer.

On va faire un exemple ensemble.
Regarde bien :
Lucie m'a dit de _____ la porte.
Clique sur continuer. Tu as quatre choix.
glodons
gloder
glodant
glode

Même si ce sont des mots inventés, tu vas voir qu'il y a en un qui fonctionne mieux avec le reste de la phrase, qui la rend correcte. Tu remarques que les quatre choix commencent par les mêmes lettres. G-L-O-D.

Quel est le meilleur choix ici? « Lucie m'a dit de gloder la porte» ressemble presque à une phrase normale, que tu pourrais entendre tous les jours ou dire toi-même. On pourrait dire « Lucie m'a dit de barrer la porte », par exemple. *Gloder*, ça ressemble à *barrer*! C'est le bon choix. Clique sur le mot *gloder*, puis sur Continuer.

Maintenant, la phrase est « Marie ___ le ciel pour voir une étoile filante ». Clique sur Continuer pour voir les choix. Il faut choisir la meilleure réponse. Tu remarques que les quatre choix commencent par les mêmes lettres. F-U-P-R.

fuprons
fuprant
fuprer
fupre

(Interaction avec les élèves) Qu'est-ce que tu en penses? Si tu penses que c'est le dernier mot qui est le bon, le mot *fupre*, tu as raison. *Fupre*, c'est le mot qui rend la phrase bien construite. On pourrait dire « Marie regarde le ciel pour voir une étoile filante ». Clique sur *Fupre*, puis sur Continuer.

La prochaine phrase est « Nous ____ adopter un bébé chat». Clique sur Continuer pour voir les choix.
druchons
drucher
druchant
druche

(Interaction avec les élèves) As-tu trouvé le bon mot? Si tu as choisi le mot *druchons*, tu as raison! C'est la fin du mot qui te fait choisir *druchons*, puisque le début est le même. On pourrait dire « Nous voulons adopter un bébé chat». Clique sur *druchons*, puis sur Continuer.

La dernière phrase est « Il a aperçu un aigle en ____ dans la forêt». Clique sur Continuer pour voir les choix.
coubrer
coubrant
coubrons
coubre

Quelle est la meilleure réponse? Si tu as choisi le mot *coubrant*, tu as raison! C'est la fin du mot qui te fait choisir *coubrant*. On pourrait dire « Il a aperçu un aigle en marchant dans la forêt».

Tu es maintenant prêt ou prête pour l'Épreuve! Clique sur Épreuve. Essaie de répondre vite et bien.

ANNEXE B

PROTOCOLE DE PASSATION, ÉPREUVES D'HABILETÉS DOMAIN-GENERAL / NON VERBALES

Sous-test 6 CFT1-R : **Matrices**, 2 minutes ½ (Cattell *et al.*, 1973a), 30 secondes par item (Wechsler, 2016)

Voir fichier Epreuves.psyexp

Consignes orales

Dans ce nouvel exercice, tu vas voir apparaitre plusieurs figures. En haut, il y a une matrice à quatre cases, dans laquelle il manque une figure. Trois des cases contiennent des dessins, mais le dessin en bas à droite est manquant. Tu devras compléter la matrice en cliquant sur une des figures présentées en dessous. Tu choisis le dessin qui va dans l'espace vide.

On va faire trois exemples ensemble.

Regarde ces images. Laquelle parmi celles-ci (pointer les possibilités de réponses) va ici (pointer la case avec le point d'interrogation) ? C'est cela (pointer la réponse) *le triangle blanc situé en bas à droite dans le carré* est la seule réponse qui va en ligne et en colonne.

Tu auras pour chaque item 30 secondes pour choisir la bonne image. Quand il te sera demandé de faire une pause, attends. On attendra que tout le monde finisse avant de passer au deuxième exercice.

Consignes dans le logiciel

« Bienvenue!

Cette épreuve est composée de deux exercices. Clique pour passer aux consignes du premier exercice.

Dans l'exercice qui suit, un grand carré contenant quatres petites cases se trouvera au centre. Trois de ces cases contiennent des dessins, mais le dessin en bas à droite est manquant. L'une des 5 cases présentées sous le grand carré correspond correctement à la case vide. Tu dois cliquer sur la bonne réponse parmi ces 5 cases.

Clique pour passer à la pratique.

Exemple :

Clique sur le dessin qui va dans l'espace vide.

Tu peux prendre une pause.

Quand tu seras prêt ou prête, clique pour passer à l'exercice.

C'est le temps de prendre une pause!

Quand on te l'indiquera, clique pour passer à l'explication du deuxième exercice. »

Sous-test 5 CFT1-R : **Classifications**, 4 minutes (Cattell *et al.*, 1973a)

Consignes orales

Dans ce nouvel exercice, tu vas voir apparaitre 5 figures. Tu dois sélectionner celle qui est différente des autres.

Quatre des figures dans chaque exemple ont des formes qui se ressemblent d'une façon ou d'une autre, mais une figure est différente des quatre autres. Pour chaque rangée, tu dois trouver la figure qui est différente des autres. Lorsque tu la trouves, clique dessus.

Consignes dans le logiciel

« Cinq figures apparaitront à l'écran. Quatre des figures ont des formes qui se ressemblent d'une façon ou d'une autre, mais une figure est différente des quatre autres. Tu dois trouver la figure qui est différente des autres. Lorsque tu la trouves, clique dessus.

Clique pour passer à la pratique.

Exemple :

Tu peux prendre une pause.

Quand tu seras prêt ou prête, clique pour passer à l'exercice.

Choisis la figure qui est différente. »

ANNEXE C

PROTOCOLE DE PASSATION, ÉPREUVES DE COMPRÉHENSION

Épreuve de compréhension en lecture avec courts textes

Consignes orales :

Maintenant, tu vas lire plusieurs petits textes. Tu devras ensuite de déterminer si les énoncés présentés à la suite de chaque texte **représentent**, ou **se rapporten**t à l'histoire que lue juste avant. Tu ne peux pas retourner au texte après avoir cliqué sur Continuer. Lis suffisamment le texte et que tu auras l'impression de l'avoir compris, tu pourras cliquer sur Continuer. 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 tous les textes.

Quand tu es prêt ou prête, clique sur Commencer. Essaie de répondre vite et bien.

Consignes dans le logiciel :

« Tu vas lire en tout 10 petits textes. Ces textes seront chaque fois suivis de 4 énoncés. Tu devras décider si OUI ou NON, ces énoncés représentent, ou se rapportent à l'histoire lue, en appuyant sur la touche O, pour oui, ou sur la touche N, pour non.

Nous ferons d'abord un exemple. Après avoir lu le texte qui suit, tu devras appuyer sur une touche du clavier pour voir les énoncés.

Lis à ta vitesse, pour bien comprendre le texte.

Pour commencer, appuie sur une touche du clavier. »

Épreuve de compréhension en lecture avec long texte

Les instructions sont écrites (voir annexe G).

ANNEXE D

PROTOCOLE DE PASSATION, ÉPREUVES DE CONSCIENCE SYNTAXIQUE

Consignes orales générales pour les épreuves de conscience syntaxique

On va faire des exercices qui vont porter sur des phrases avec erreurs, qu'on va corriger. On va aussi reproduire des erreurs, mais attention, des erreurs contrôlées ! Ce qui rend l'exercice un peu plus difficile, mais intéressant aussi…

Ici, ce qu'on veut faire, c'est découvrir comment la langue fonctionne, et réaliser qu'on peut jouer avec la langue, aussi. Comme avec des blocs légo, on *déplace* des blocs de mots, on en *ajoute*, on en *enlève*, on les *remplace* par d'autres…
Dans cet exercice, les premières phrases de chaque question sont fausses : elles comportent une erreur. Les deux phrases qui suivent ne comportent pas d'erreur. Pour chaque numéro ou bloc, tu dois faire quatre choses :

1. Identifier l'erreur dans une phrase mal construite
2. Corriger cette erreur
3. Refaire ou reproduire l'erreur de la première phrase dans une deuxième phrase.
4. Refaire ou reproduire l'erreur de la première phrase dans une troisième phrase.

Il y aura chaque fois deux blocs d'entrainement, ou de pratique, avant de commencer les exercices, qui sont au nombre de 3.
L'erreur est toujours dans la structure, pas dans la graphie ! (ex. les majuscules et les points ne font pas partie de ce qui cloche). Il ne faut pas passer trop de temps sur chaque bloc. Tu essaies juste de faire de ton mieux, puis tu passes au bloc suivant.

C'est parti !

Remplacement

Consignes dans le logiciel :

Consignes :

Cet exercice fonctionne par bloc de trois phrases, en quatre temps.

Premièrement, tu verras une phrase qui comporte une erreur. Tu vas devoir cliquer sur le mot qui pose problème.

Deuxièmement, tu vas devoir rendre la phrase correcte en écrivant un nouveau mot.

Troisièmement, une nouvelle phrase est présentée à l'écran. Elle est bien écrite, sans erreur. Ton rôle est de reproduire l'erreur de la première phrase dans cette deuxième phrase.

Quatrièmement, tu dois aussi reproduire cette même erreur dans la troisième phrase, qui apparaitra ensuite.

Lorsqu'on t'en aura donné l'autorisation, clique sur « Continuer ».

Consignes orales :

1er exemple

le chat dont ta sœur aime est doux.
la tortue que vous observez est vieille.
le coiffeur bavard que ma mère adore utilise des ciseaux coupants.

Qu'est-ce qu'on fait pour faire fonctionner la première phrase ? Ce qui cloche, c'est le subordonnant dont. On clique dessus.

On doit changer le subordonnant dont pour que. le chat que ta sœur aime est doux.

Donc si on veut reproduire l'erreur sur la phrase suivante, il faut d'abord identifier le subordonnant.

la tortue que vous observez est vieille. On remet le dont, pour rendre la phrase incorrecte comme celle du début. la tortue dont vous observez est vieille.

le coiffeur bavard que ma mère adore utilise des ciseaux coupants. On remet le dont, pour rendre la phrase incorrecte comme celle du début.

2e exemple

mon chien va par le vétérinaire
sa mère parle de son anniversaire
le garçon turbulent obéit à son professeur dévoué

Qu'est-ce qu'on fait pour faire fonctionner la première phrase ? Ce qui cloche, c'est la préposition par. On clique dessus.

Ici, la préposition doit être remplacée : mon chien va chez le vétérinaire.

Si on veut reproduire l'erreur sur la phrase suivante, il faut d'abord identifier la préposition.

sa mère parle de son anniversaire. On remplace par « par » pour reproduire l'erreur.

le garçon turbulent obéit à son professeur dévoué. On remplace par « par » pour reproduire l'erreur.

Addition suppression

Consignes dans le logiciel :

Cet exercice fonctionne par bloc de trois phrases, en quatre temps.

Premièrement, tu verras une phrase qui comporte une erreur. Il manque un mot ou un groupe de mots dedans. Tu vas devoir cliquer sur le point dans la phrase où il manque un mot ou un groupe de mots.

Deuxièmement, tu vas devoir rendre la phrase correcte en écrivant un nouveau mot ou un nouveau groupe de mots.

Troisièmement, une nouvelle phrase est présentée à l'écran. Elle est bien écrite, sans erreur. Ton rôle est de cliquer sur le mot ou le groupe de mots à enlever pour reproduire l'erreur de la première phrase dans cette deuxième phrase.

Quatrièmement, tu vas devoir aussi reproduire cette même erreur dans la troisième phrase, qui apparaitra ensuite.

Lorsqu'on t'en aura donné l'autorisation, clique sur « Continuer ».

Clique sur le point dans la phrase où il manque un mot ou un groupe de mots, puis clique sur « Continuer ».

Clique sur le mot ou le groupe de mots à enlever pour reproduire l'erreur de la phrase du début dans la phrase suivante, puis clique sur « Continuer ».

Clique sur le mot ou le groupe de mots à enlever pour reproduire l'erreur de la phrase du début dans la phrase suivante, puis clique sur « Continuer ».

Consignes orale :

1er exemple

le copain ma mère est gentil.
le chien de Louis est doux.
ces chats roux du Portugal semblent très fatigués.

Qu'est-ce qu'on fait pour faire fonctionner la première phrase ?

Le copain de ma mère est gentil. Il faut ajouter la préposition de, parce qu'il avait été supprimé.

Donc si on veut reproduire l'erreur sur la phrase suivante, il faut d'abord identifier la préposition.

Le chien de Louis est doux. On enlève le « de » pour reproduire l'erreur.

ces chats roux du Portugal semblent très fatigués. Ou est la préposition ici ? du ! On enlève le « du » pour reproduire l'erreur.

2e exemple

le lézard recherche __ .
le magicien lève le bras.
une télévision bruyante diffuse des images choquantes.

Qu'est-ce qu'on fait pour faire fonctionner la première phrase ? Alors là il manque plusieurs mots, pas juste un.

Ici, on doit compléter la phrase avec un complément : le lézard recherche __ quelque chose, la pomme, son amie, etc.

Si on veut reproduire l'erreur sur la phrase suivante, il faut d'abord identifier le complément direct.

le magicien lève le bras. On doit enlever le complément « le bras » pour reproduire l'erreur.

une télévision bruyante diffuse des images choquantes. On doit enlever le complément direct des images choquantes pour reproduire l'erreur.

Il faut chercher ce qui joue le même rôle dans la phrase.

Déplacement

Consignes dans le logiciel :

Cet exercice fonctionne par bloc de trois phrases, en quatre temps.

Premièrement, tu verras une phrase qui comporte une erreur. Un mot ou un groupe de mots **est mal placé**. Tu vas devoir cliquer sur le mot ou le groupe de mots qui pose problème.

Deuxièmement, tu vas devoir rendre la phrase correcte en cliquant sur l'endroit dans la phrase où déplacer ce mot ou ce groupe de mots.

Troisièmement, une nouvelle phrase est présentée à l'écran. Elle est bien écrite, sans erreur. Ton rôle est de cliquer sur le mot ou le groupe de mots à déplacer pour reproduire l'erreur de la première phrase dans cette deuxième phrase, puis sur l'endroit où le déplacer.

Quatrièmement, tu vas devoir aussi reproduire cette même erreur dans la troisième phrase, qui apparaitra ensuite.

Lorsqu'on t'en aura donné l'autorisation, clique sur « Continuer ».

Consignes orales :

1er exemple

des fleurs la femme apporte.
le fermier loue des tracteurs.
un petit robot répare la voiture brisée.

Qu'est-ce qu'on fait pour faire fonctionner la première phrase ?

La femme apporte quoi ? des fleurs ! Quel rôle il joue dans la phrase, ce groupe de mots-là, « des fleurs ». C'est un complément direct.

Pour faire fonctionner la phrase, on déplace le groupe « des fleurs ».

Donc si on veut reproduire l'erreur sur la phrase suivante, il faut d'abord identifier le complément direct.

Le fermier loue quoi ? des tracteurs ! On déplace « des tracteurs » au début de la phrase pour reproduire l'erreur.

Un petit robot répare quoi ? la voiture brisée ! On déplace « la voiture brisée » au début de la phrase pour reproduire l'erreur.

2e exemple

les enfants le texte récitent
la femme mange la pomme
les enfants cueillent les fleurs sauvages

Qu'est-ce qu'on fait pour faire fonctionner la première phrase ?

Ici, le verbe doit être placé au bon endroit : les enfants *récitent* le texte.

Si on veut reproduire l'erreur sur la phrase suivante, il faut d'abord identifier le verbe.

La femme mange la pomme ! On déplace « mange » à la fin de la phrase pour reproduire l'erreur.

Les enfants cueillent les fleurs sauvages. On déplace « cueillent » à la fin de la phrase pour reproduire l'erreur.

ANNEXE E

ITEMS DES ÉPREUVES DE CONSCIENCE SYNTAXIQUE

Numéro d'item	Items	Phrase	Opération	Catégorie	Unité	Fonction/rôle de l'élément erroné
P ent 1	le chat dont ta sœur aime est doux.	modèle	remplacement	Sub	mot	Subordonnant dans Psub CdN
P ent 1a	la tortue que vous observez est vieille.	cibleA	remplacement	Sub	mot	Subordonnant dans Psub CdN
P ent 1b	le coiffeur bavard que ma mère adore utilise des ciseaux coupants.	cibleB	remplacement	Sub	mot	Subordonnant dans Psub CdN
P ent 2	mon chien va par le vétérinaire	modèle	remplacement	Prép	mot	Prép dans CI du GV
P ent 2a	sa mère parle de son anniversaire	cibleA	remplacement	Prép	mot	Prép dans CI du GV
P ent 2b	le garçon turbulent obéit à son professeur dévoué	cibleB	remplacement	Prép	mot	Prép dans CI du GV
11	le linge sèche pour le jardin.	modèle	remplacement	Prép	mot	Prép dans CI du GV
11A	la tortue nage dans la mer.	cibleA	remplacement	Prép	mot	Prép dans CI du GV

11B	toutes ces petites embarcations avancent vers le rivage.	cibleB	remplacement	Prép	mot	Prép dans CI du GV
12	les gâteaux dont ta sœur veut manger sont roses.	modèle	remplacement	Sub	mot	Subordonnant dans Psub CdN
12A	les fruits que tu as cueillis sont mûrs.	cibleA	remplacement	Sub	mot	Subordonnant dans Psub CdN
12B	ce petit enfant qui a de jolies joues rouges regarde ton frère	cibleB	remplacement	Sub	mot	Subordonnant dans Psub CdN
13	le marié a choisi une bague qui la couleur est très belle.	modèle	remplacement	Sub	mot	Subordonnant dans Psub CdN
13A	le musicien a chanté une chanson dont la mélodie est ancienne.	cibleA	remplacement	Sub	mot	Subordonnant dans Psub CdN
13B	nous avons vu cet ennuyeux spectacle que mon grand frère a détesté.	cibleB	remplacement	Sub	mot	Subordonnant dans Psub CdN
14	mon amie commence pour le gâteau	modèle	remplacement	Prép	mot	Prép dans CI du GV
14A	sa mère doute de son honnêteté	cibleA	remplacement	Prép	mot	Prép dans CI du GV
14B	ils réfléchissent au spectacle	cibleB	remplacement	Prép	mot	Prép dans CI du GV
P ent 3	le copain ___ ma mère est gentil.	modèle	suppression	Prép	mot	Prép dans CdN
P ent 3a	le chien de Louis est doux.	cibleA	suppression	Prép	mot	Prép dans CdN
P ent 3b	ces chats roux du Portugal semblent très fatigués.	cibleB	suppression	Prép	mot	Prép dans CdN
P ent 4	Le lézard recherche __ .	modèle	suppression	Dét	mot	Dét dans SN CD
P ent 4a	Le magicien lève le bras.	cibleA	suppression	Dét	mot	Dét dans SN CD

P ent 4b	Une télévision bruyante diffuse des images choquantes.	cibleB	suppression	Dét	mot	Dét dans SN CD
2	le garçon attrape ___.	modèle	suppression	GN	syntagme	CD
2A	le médecin offre des conseils.	cibleA	suppression	GN	syntagme	CD
2B	ce courageux chevalier poursuit le gros dragon.	cibleB	suppression	GN	syntagme	CD
10	___ visitent le musée.	modèle	suppression	GN	syntagme	Sujet
10A	les bateaux quittent le port.	cibleA	suppression	GN	syntagme	Sujet
10B	tous les jeunes joueurs fêtent la victoire de leur équipe.	cibleB	suppression	GN	syntagme	Sujet
16	le héros combat __.	modèle	suppression	GN	syntagme	CD
16A	la chanteuse apprécie son public.	cibleA	suppression	GN	syntagme	CD
16B	ces joueurs mécontents contestent la décision injuste.	cibleB	suppression	GN	syntagme	CD
17	le lion __ le zèbre.	modèle	suppression	V	mot	V
17A	le garçon cueille la rose.	cibleA	suppression	V	mot	V
17B	elle adore son nouveau jouet.	cibleB	suppression	V	mot	V
19	le lapin __ mange sa laitue s'est blessé	modèle	suppression	Sub	mot	Subordonnant dans Psub CdN
19A	le pantalon que la fille porte est déchiré	cibleA	suppression	Sub	mot	Subordonnant dans Psub CdN
19B	cette plante verte dont je t'ai parlé est délicate	cibleB	suppression	Sub	mot	Subordonnant dans Psub CdN
21	il accepte le bisou ___ sa tante.	modèle	suppression	Prép	mot	Prép dans CI du verbe
21A	elle colle les étiquettes sur le mur.	cibleA	suppression	Prép	mot	Prép dans CI du verbe

21B	le professeur fatigué écrit des bêtises au tableau.	cibleB	suppression	Prép	mot	Prép dans CI du verbe
P ent 2	des fleurs la femme apporte.	modèle	déplacement	GN	syntagme	CD
P ent 2a	le fermier loue des tracteurs.	cibleA	déplacement	GN	syntagme	CD
P ent 2b	un petit robot répare la voiture brisée.	cibleB	déplacement	GN	syntagme	CD
P ent 2	les enfants le texte récitent	modèle	déplacement	V	mot	V
P ent 2a	la femme mange la pomme	cibleA	déplacement	V	mot	V
P ent 2b	les enfants cueillent les fleurs sauvages	cibleB	déplacement	V	mot	V
1	lave la vaisselle le papa.	modèle	déplacement	GN	syntagme	Sujet
1A	le garçon regarde un film.	cibleA	déplacement	GN	syntagme	Sujet
1B	ce fier artisan fabrique un fromage délicieux.	cibleB	déplacement	GN	syntagme	Sujet
4	son frère suit de natation des cours.	modèle	déplacement	Gprép	syntagme	CdN
4A	cet ami veut une tasse de café	cibleA	déplacement	Gprép	syntagme	CdN
4B	le jeune garçon porte une veste de laine jaune	cibleB	déplacement	Gprép	syntagme	CdN
5	sa sœur les films de science-fiction aime.	modèle	déplacement	V	mot	V
5A	sa comptable adore les pièces de monnaie.	cibleA	déplacement	V	mot	V
5B	le bébé gourmand mange le délicieux pain aux olives.	cibleB	déplacement	V	mot	V
6	Adélie les cheveux doux a.	modèle	déplacement	V	mot	V
6A	Vincent termine son devoir.	cibleA	déplacement	V	mot	V
6B	les jeunes chiots mordent les grosses fleurs	cibleB	déplacement	V	mot	V

9	un chapeau le singe porte.	modèle	déplacement	GN	syntagme	CD
9A	le bébé dévore un biscuit.	cibleA	déplacement	GN	syntagme	CD
9B	le randonneur fatigué abandonne le sentier difficile	cibleB	déplacement	GN	syntagme	CD
15	fleurie elle porte une jupe.	modèle	déplacement	GAdj	syntagme	CdN
15A	il aime une fille gentille.	cibleA	déplacement	GAdj	syntagme	CdN
15B	les citrons verts ont une saveur sucrée.	cibleB	déplacement	GAdj	syntagme	CdN
18	le renard est gentil qui joue dans la cage.	modèle	déplacement	Psub	syntagme	CdN
18A	la maison où vivent ses grands-parents parait énorme.	cibleA	déplacement	Psub	syntagme	CdN
18B	ce gros chat gris que ta sœur adore semble toujours affamé.	cibleB	déplacement	Psub	syntagme	CdN
20	survole le lac le colibri.	modèle	déplacement	GN	syntagme	Sujet
20A	le zèbre traverse la savane.	cibleA	déplacement	GN	syntagme	Sujet
20B	une jeune chèvre bondit dans les prés verts.	cibleB	déplacement	GN	syntagme	Sujet

ANNEXE F

ITEMS DES ÉPREUVES DE CONSCIENCE MORPHOLOGIQUE

Épreuve de paires de mots ou de jugement de relations de mots

Entrainement :

1. école, écolier
2. heure, heureux
3. dent, dentiste
4. bague, baguette

Items de l'épreuve :

1. glisser, glissade saler, salade; V>N; F58; AF42; F54; F52	'Action de X'
2. baigner, baignade couler, couleur V>N; F53; F45; F58; F65	'Action de X'
3. livrer, livraison ouvrir, ouvrier V>N F63; F55; TF71; F59	'Action de X'
4. comparer, comparaison pousser, poussière V>N F62; F58; F64; F54	'Action de X'
5. danser, danseur sentir, sentier V>N; F57; F55; F67; F54	'Celui qui X'
6. chasser, chasseur saisir, saison V>N ; F60; F56; F62; F66	'Celui qui X'
7. coffre, coffret gaz, gazon N>N; F52; F49; F62; F49	'Petit X'
8. livre, livret balle ballet N>N; F53; F53; F59; F51	'Petit X'
9. feuille, feuillage cour, courage; N>N; F60; F46; F67; F61	'Ensemble de X'
10. ville, village or, orage N>N; TF72; F64; F64; F52	
11. guerre, guerrier chant, chantier N>N; F69; F51; F57; F59	'Lieu'
12. couture, couturier colle, collier N>N; F50; F48; F47; F49	'Celui qui X'
	'Celui qui

X/marchand’

Légende **:** AF = assez fréquent, F = fréquent, TF = très fréquent. Le chiffre indiqué à côté de la classe de fréquences correspond à l’indice de fréquence selon la nomenclature du logiciel Antidote ©, version 9.2.

Épreuve d’extraction de la base

Entrainement

1. tassette, tasse (base+affixe = pseudo-mot)
2. alouette, alouette (base+pseudo-affixe =mot)
3. refaire, faire (base+affixe = mot)
4. renifler, renifler (base+pseudo-affixe =mot)
5. patinage, patine (base+affixe = mot)
6. partirage, partir (base+affixe = pseudo-mot)

Items de l’épreuve :

	Base+affixe (=pseudo-mot), sur le modèle dégarer ou tassette	Base+affixe (=mot), sur le modèle fillette ou refaire	Base+pseudo-affixe, sur le modèle alouette ou renifler
Préfixé	dénager – V emmanger – V replaire – V prédormir – V	défaire – V (F56) emporter – V (F61) repartir – V (F58) précuire – V (AF32)	décider –V (F69) employer – V (F63) reprocher – V (F60) préférer – V (F66)
Suffixé	bruitet – N lampier – N finirage – N dérangeur – N	garçonnet – N (AF37) banquier – N (F54) lavage – N (F48) utilisateur – N (F60)	chalet – N (F49) papier – N (F63) voyage – N (F65) odeur – N (F56)

Épreuve de dérivation

Entrainement :

1. garage. Gaston est un garagiste.

2. délice. Le repas est délicieux.
3. conducteur. Adélie est une bonne conductrice.
4. parler. Nous parlons.

Items de l'épreuve :

1. efficace. Elle est fière de son efficacité.
2. simple. On remarque sa simplicité.
3. riche. C'était un roi d'une fabuleuse richesse.
4. vite. Cette voiture est remarquée pour sa vitesse.
5. ferme. Bill est un fermier.
6. aventure. Émile est un aventurier.
7. maire. Catherine travaille à la mairie.
8. boucher. Papa doit passer à la boucherie.
9. mystère. Cet homme à lunettes a un air mystérieux.
10. silence. Nous sommes dans un endroit silencieux.
11. barbe. Cet homme est vraiment barbu.
12. pointe. Ce crayon est bien pointu.
13. remarquer. Son talent en dessin est remarquable.
14. adorer. Paméla est une fille adorable.
15. décorer. Maman utilise son ruban décoratif.
16. créer. Milan est un garçon très créatif.
17. guerrier. Athéna est une redoutable guerrière.
18. prince.Il faut parler à la princesse.
19. sérieux. Catherine est une fille.
20. pensif. Maman est très pensive.
21. jouer. Nous jouons.
22. dormir. Nous dormons.
23. passer. Nous passons.
24. écouter. Nous écoutons.

Épreuve de décomposition

Entrainement :

1. chanteur. Mon grand frère adore chanter.
2. douceur. Le poil du chien est très doux.
3. animatrice. Gaston est un bon animateur.
4. chassons. Il faut chasser.

Items de l'épreuve :

1. débrouillardise. Luigi est très débrouillard.
2. gourmandise. C'est un bébé très gourmand.

3. lenteur. La tortue est un animal très lent.
4. grosseur. Mon lapin est trop gros.
5. courageux. Mario a beaucoup de courage.
6. chanceux. Albert a eu de la chance.
7. dramatique. La coupe du gros arbre a été un drame.
8. climatique. La région est connue pour son drôle de climat.
9. voisinage. Marcus est mon voisin.
10. ombrage. L'arbre fait de l'ombre.
11. moustiquaire. Renée a aperçu un gros moustique.
12. bibliothécaire. Patricia adore aller à la bibliothèque.
13. travaillant. Bob aime beaucoup travailler.
14. bouillant. La soupe est en train de bouillir.
15. pensif. Denis va sûrement y penser.
16. inventif. Cet ami a du talent pour inventer.
17. chanteuse. Bastien est un excellent chanteur.
18. fatigante. Mon chien est un peu fatigant.
19. active. Mon oncle est très actif.
20. horlogère. Cet homme est horloger.
21. jouons. Il faut jouer.
22. dormons. Il faut dormir.
23. créons. Il faut créer.
24. lisons. Il faut lire.

Épreuve de choix d'affixe dérivationnels dans des pseudomots

Entrainement :

1. J'aimerais que mon chaton arrête de _____ .
   bluvable
   bluvage
   bluvir*
2. Mon petit frère est de plus en plus _____ .
   prilusage
   prilusant*
   prilusir
3. Hubert et son papa vont à la _____ .
   louviterie*
   louvitir
   louvitable

Items de l'épreuve :

1. Amélie a peur de _____ .

tanvage
tanvir*
tanvable
2. Martine adore _____ avec sa poupée.
nouvant
nouvir*
nouvade
3. Léa aimerait apprendre à _____ .
glomir*
glomeau
glomal
4. Les pompiers ont l'habitude de _____ .
poglure
poglesque
poglir*
5. La princesse n'aime pas _____.
pocrir*
pocrible
pocrité
6. J'ai vu un _____ bébé chien.
brotable*
brotité
brotir
7. Ce film est _____.
domprir
domprant*
domprure
8. Leslie a donné un surnom _____ à son chien.
jouvir
jouveau
jouval*
9. La maison du géant est _____ .
stochir
stochade
stochesque*
10. Le chevalier est très _____ .
pachible*
pachir
pachage
11. Le _____ permet d'avoir la peau lisse.
banflal
banflir
banflage*
12. Cette _____ est délicieuse.

tambresque
tambrade*
tambrir
13. Paco est un petit _____.
doubible
doubeau*
doubir
14. Elle travaille très fort sur sa _____.
dovir
dovable
dovure*
15. La _____ des dauphins est bien connue.
joupité*
joupante
joupir

Épreuve de choix d'affixe flexionnels dans des pseudomots

Entrainement :

1. Lucie m'a dit de _____ la porte.
glodons
gloder*
glodant
glode
2. Marie _____ le ciel pour voir une étoile filante.
fuprons
fuprant
fuprer
fupre*
3. Nous _____ adopter un bébé chat.
druchons*
drucher
druchant
druche
4. Il a aperçu un aigle en _____ dans la forêt.
coubrer
coubrant*
coubrons
coubre

Items de l'épreuve :

1. L’enseignante lui a montré comment _____ de longs mots.
sidons
sider*
sidant
side
2. Mario _____ les journaux pour trouver un logement.
somprons
somprant
somprer
sompre*
3. Mes parents doivent _____ une nouvelle voiture.
monflons
monfler*
monflant
monfle
4. Nous _____ voir un bon film.
sarchons*
sarcher
sarchant
sarche
5. Elle a vu un animal bizarre en _____ la rivière.
tambrer
tambrant*
tambrons
tambre
6. Ils ont des employés _____ avec passion.
banflant*
banfler
banflons
banfle
7. Les familles _____ un chien sont bienvenues dans le parc.
torant*
tore
torons
torer
8. Nous _____ des petites tortues de mer qui pondent la nuit.
pache
pacher
pachant
pachons*
9. J’aimerais vous _____ à une fête d’anniversaire.
pidons
pider*
pidant

pide
10. Nous _____ l'actrice qui joue dans ce film.
vorons*
vorer
vorant
vore
11. Jules a vu une étoile filante en _____ attentivement le ciel.
tutons
tuter
tutant*
tute
12. Le cordonnier _____ des trous dans la ceinture avec un poinçon.
drachons
dracher
drachant
drache*
13. Il continue à _____ malgré la fatigue.
dinuons
dinuer*
dinuant
dinue
14. Le jeune loup _____ les plus vieux à traverser la forêt.
trupons
truper
trupant
trupe*
15. La talentueuse étudiante _____ très bien en parlant.
peutrons
peutrer
peutrant
peutre*
16. Avant de partir de la maison, nous _____ la vaisselle.
churrons*
churrer
churrant
churre

ANNEXE G

ITEMS DES DEUX ÉPREUVES DE COMPRÉHENSION DE TEXTES

Épreuve de compréhension en lecture, textes courts, réponses à choix, sur ordinateur

bloc	textes	Rep1		Rep2		Rep3		Rep4	
1	Il fait beau, ce matin, sur les côtes de Mumbai. C'est marée basse, tout semble calme dans cette petite mare. Oh non ! : un bébé poisson-ballon est resté prisonnier de la petite mare formée par la mer en se retirant. Le petit poisson-ballon nage vers la gauche, puis vers la droite, et se cogne chaque fois contre un rocher. « Au secours ! hurle-t-il, je ne peux pas sortir ! » Pof ! Il se gonfle comme un ballon pour se protéger du danger. Une pieuvre majestueuse, qui était cachée sous un rocher, se glisse soudain dans la mare. Intimidé, le petit poisson-ballon la fixe en silence. « Tu crois que tu es prisonnier de la mare, explique la pieuvre, mais la marée remontera bientôt. Patience ! Tu pourras bientôt rentrer chez toi. » Le bébé poisson-ballon est soulagé. C'est alors que l'eau s'engouffre dans la mare. Le niveau de l'eau monte. Le petit poisson-ballon s'élève avec l'eau et passe par-dessus les rochers. Il est fou de joie. « Au revoir, petit poisson. Sois prudent ! »	C'est l'histoire d'un bébé poisson paniqué à l'idée de ne pas retrouver la mer.	3	C'est l'histoire d'un bébé poisson dans une mare, à marée basse.	1	C'est l'histoire d'un bébé poisson gros comme un ballon.	0	C'est l'histoire d'un bébé poisson protégé par les rochers.	0

2	Adil, le crocodile triste, pleurait au bord du lac. — Pourquoi pleures-tu ? demanda Raymond le papillon. — Je me sens très seul, répondit Adil, le crocodile. Je n'ai pas d'amis. J'ai essayé de devenir ami avec les humains qui ont des filets... Mais quand j'ai voulu leur dire bonjour, ils se sont enfuis en courant ! — Tu peux devenir mon ami, proposa Raymond. À condition de ne pas me manger ! — Promis, répondit Adil le crocodile. Raymond revint bientôt accompagné de tous ses amis papillons ! Adil, le crocodile, pleurait désormais de bonheur. Il n'était plus seul.	C'est l'histoire d'un crocodile.	1	C'est l'histoire d'humains gentils.	0	C'est l'histoire d'Adil, finalement heureux d'être bien entouré.	3	C'est l'histoire d'une chasse aux papillons.	0

3	En ouvrant le lave-linge, Tim n'imagine pas être confronté à l'une des plus palpitantes énigmes des temps modernes. Du haut de ses douze ans, il n'est pas très fort en matière de ménage. Lui, sa passion, c'est la science. Il se contente donc d'accumuler le linge sale dans un coin de sa chambre. Ce matin-là, Tim s'aperçoit qu'il n'a plus aucune chaussette propre dans son placard. Ses parents déjà partis au travail, il doit se débrouiller tout seul. Il saisit aussitôt masque et tuba et plonge dans son tas de linge sale à la recherche de chaussettes encore vivantes. Cinq paires sont remontées à la surface. Un peu raidies par la crasse, elles ont quand même conservé leur couleur d'origine. Tim se dirige vers la machine à laver et jette les dix chaussettes dedans. C'est la première fois qu'il manipule tous ces boutons. Il ne connait pas les capacités de ces machines redoutables. Vingt minutes de tourbillon plus tard, le cycle semble terminé. Notre héros surexcité plonge sa main fébrile dans la gueule humide de la bête. Et là : incroyable mais vrai, seules huit chaussettes sont retrouvées. Malgré une fouille méticuleuse, nul ne sait ce qui est advenu des chaussettes manquantes. « Un mystère bien plus grand que celui du Big Bang ! » en rira à son retour la maman de Tim, avec un air entendu de vieux pirate.	C'est l'histoire d'un garçon qui apprend à se débrouiller.	3	C'est l'histoire de paires de chaussettes sales.	1	C'est l'histoire de la résolution du mystère qu'est une machine à laver.	0	C'est l'histoire d'un monstre à la gueule humide.	0

4	Thomas a mis ses jolis souliers neufs. Thomas a peur. Juste un peu. Mais il est content tout de même. À chaque rentrée, c'est la même chose. Le trac lui prend et lui broie les tripes ! Oui, il en a mal au ventre ! Thomas se regarde dans le miroir, remet une dernière fois de l'ordre dans ses cheveux et attrape son sac. Sur le chemin de l'école, la tension monte. Y aura-t-il encore de cet infâme gratin de chou-fleur qui sent les pieds à la cantine ? Le directeur sera-t-il aussi dur que l'année passée ? Et les punitions ! Ah là là ! Il va falloir faire attention ! Mais c'est déjà l'heure de former un rang. Par deux. En silence. Lorsque Thomas prend la parole, tout le monde est assis. C'est parti ! — Bonjour, prenez votre cahier, nous allons commencer!	C'est l'histoire d'un élève qui vit sa première rentrée scolaire.	0	C'est l'histoire du trac que vit un directeur d'école coquet la veille d'une rentrée scolaire.	0	C'est l'histoire d'une rentrée scolaire stressante.	1	C'est l'histoire des émotions vécues par les professeurs, mais aussi par les élèves, lors de la rentrée scolaire.	3

5	Dada et Dédé les poussins manchots royaux sont inséparables. Les manchots ne volent pas, ils nagent. Mais Dada et Dédé sont trop petits pour nager. Alors, ils marchent. Un jour, ils partent faire une grande promenade. Ils marchent pendant des heures, des jours, des semaines, jusqu'à ce que leur duvet se change en plumage et que leur plumage change de couleur. Ils tombent alors sur quelque chose qu'ils n'avaient jamais vu. La mer. — Qu'est-ce qu'on fait maintenant ?	C'est l'histoire de manchots qui marchent.	1	C'est l'histoire d'amis qui grandissent.	2	C'est l'histoire d'oiseaux qui volent.	0	C'est l'histoire de poussins heureux.	0
6	Dès son arrivée, Jack saute du camion de pompiers et court à l'intérieur de la maison. Il sait qu'un garçon est piégé à l'étage. Il lutte contre les flammes pour monter les escaliers et trouve le garçon, Tim, caché derrière un lit. Soudain, ils entendent un grand fracas... Les escaliers ont disparu ! Jack ouvre la fenêtre et crie pour qu'on lui apporte une échelle. Quand ils atteignent le sol, Tim sert sa mère dans ses bras, et Jack est un héros.	C'est l'histoire du sauvetage d'un enfant par un pompier.	2	C'est l'histoire d'un policier appelé Jack.	0	C'est l'histoire d'escaliers détruits par les flammes.	1	C'est l'histoire de l'efficacité de l'eau contre les incendies.	0

7	Emma n'avait jamais campé auparavant. Elle s'est réveillée dans le noir, en entendant un reniflement bizarre dehors. Au fur et à mesure qu'Emma écoutait, le bruit s'approchait. Elle a attrapé sa sœur et lui a dit qu'un énorme monstre arrivait. Quand sa sœur a regardé dehors, elle a ri, et a dit à Emma de venir voir. Ce n'était qu'un hérisson à la recherche de limaces. Emma était étonnée qu'une si petite créature puisse faire autant de bruit.	C'est l'histoire d'une fillette qui a vu une créature effrayante.	0	C'est l'histoire d'une première nuit de camping.	1	C'est l'histoire d'une peur injustifiée.	2	C'est l'histoire d'un rêve terrifiant.	0
8	Il était une fois, il y a plusieurs millions d'années, un immense lézard caché dans la forêt. Immobile comme une statue, du bout du museau à la pointe de la queue, le lézard attendait l'arrivée d'une proie. Seuls ses yeux perçants remuaient dans tous les sens. Ses narines frémirent en reconnaissant l'odeur du délicieux insecte qui allait constituer son dîner. Le lézard géant s'apprêtait à bondir sur lui, lorsque la terre trembla. Tous les animaux de la forêt se mirent tout à coup à courir dans toutes les directions, manquant d'écraser au passage notre pauvre lézard géant. Fâché, celui-ci se retourna pour voir ce qui semait cette confusion. GLOUPS ! Une bête effrayante fonçait à travers la forêt. Elle était plus grande que les plus grands arbres et sa peau paraissait plus rugueuse que la pierre. Notre chasseur d'insectes allait-il devenir son dîner à lui ?	C'est l'histoire d'un gros lézard préhistorique.	1	C'est l'histoire des repas à l'époque des dinosaures.	2	C'est l'histoire d'un lézard écrasé par des animaux fuyant le danger.	0	C'est l'histoire du dîner d'un insecte.	0

9	La légende raconte qu'un jour, il y eut un immense incendie dans la forêt amazonienne. Tous les animaux observaient le désastre, terrifiés et impuissants. Seul le petit colibri s'activait, allant chercher quelques gouttes d'eau dans son bec pour les jeter sur le feu. Au bout d'un moment, les autres animaux, agacés par ses agissements dérisoires, lui dirent : « Colibri, tu es fou ! Tu crois que c'est avec ces gouttes d'eau que tu vas éteindre le feu ? » « Non, mais je fais ma part » dit le petit colibri. Le tatou, l'aigle, le singe et les autres animaux se regardèrent, perplexes. Ils s'élancèrent alors tous vers la rivière, prirent de l'eau dans leur gueule, leurs pattes ou leur bec et arrosèrent les arbres, les feuilles, et la terre, puis recommencèrent, inlassablement. Ces milliers de gouttes d'eau formèrent une pluie si fine et si dense que le feu finit par s'éteindre. Depuis ce jour, les arbres verdissent et l'harmonie est revenue en Amazonie.	C'est l'histoire d'une légende amérindienne.	2	C'est l'histoire de la paresse des animaux terrestres.	0	C'est l'histoire d'un incendie en Amazonie.	1	C'est l'histoire de la naissance d'un mouvement de groupe.	3
10	Au commencement, il y avait sur Terre une seule femme et un seul homme. Un jour, ils eurent à faire un choix difficile : mourir comme la lune ou comme un bananier. La lune reste invisible tous les mois pendant quelques jours, et réapparaît ensuite. Elle continue à exister par elle-même, bien qu'elle semble mourir tous les mois et revivre après quelques jours. Le bananier, lui, meurt complètement, après avoir donné ses fruits. L'homme et la femme réfléchirent un peu, puis firent un souhait : « Nous préférons mourir comme le bananier, qui laisse après lui des successeurs. » C'est depuis lors, dit-on, que les humains meurent, en laissant sur terre des enfants pour les remplacer.	C'est l'histoire de bananes délicieuses.	0	C'est l'histoire de différents cycles	3	C'est l'histoire du cycle du soleil	0	C'est l'histoire d'une femme et d'un homme.	1

Épreuve de compréhension en lecture, texte long, sur papier

CONSIGNES

Lis le texte suivant, à ta vitesse.
Des questions de compréhension te seront posées après la lecture.
Quand tu auras fini de lire le texte, lève la main. On te donnera une fiche avec des questions de compréhension sur deux pages recto-verso.
Tu ne pourras pas retourner au texte après avoir commencé à répondre aux questions.

Texte

Il était une fois un homme qui était très difficile à satisfaire. Il pensait que sa femme ne faisait rien de bien dans la maison et qu'elle était paresseuse. Lui, il travaillait dur dans les champs, chaque jour ! En supportant la chaleur, les maux de dos et les moustiques ! Un soir, pendant la saison des récoltes, il rentra à la maison, grondant et jurant, comme d'habitude.

Sa femme lui dit : « J'aimerais que tu ne sois pas toujours en colère contre ma façon de prendre soin de la maison. J'essaie de faire de mon mieux dans ce travail. » Elle eut alors une idée et lui lança un défi : « Demain, échangeons nos rôles. Je vais sortir avec les moissonneurs pour récolter le foin, et tu t'occuperas de la maison. » Le mari pensa que ce pouvait être une bonne idée. Il était certain que prendre soin de la maison était plus simple que d'aller aux champs.

Ainsi, tôt le lendemain matin, sa femme prit la faux et sortit dans le champ de foin avec le reste des moissonneurs. Entre-temps, son mari devait s'occuper de la maison et faire le travail qu'elle faisait habituellement. Bien entendu, il devait aussi le faire beaucoup mieux et plus à son goût à lui !

La première chose qu'il avait besoin de faire était de baratter du beurre. Après avoir baratté pendant un certain temps, il eut soif. Il descendit donc à la cave pour prendre

une chope de bière. Alors qu'il plaçait le verre sous le robinet ouvert, il entendit le cochon marcher au-dessus de sa tête, dans la cuisine. Il monta à toute vitesse les escaliers pour s'assurer que le cochon ne gâcherait pas le barattage. Arrivé à l'étage, il vit que le cochon avait déjà renversé le récipient et qu'il se roulait et grognait joyeusement dans la crème. Dans sa fureur, l'homme oublia tout ce qu'il avait commencé en bas et se mit à hurler et à poursuivre le cochon dans la maison. Le cochon eut si peur qu'il sortit de la maison et se sauva à toute allure.

L'homme se souvint tout à coup du tonneau de bière. À la cave, il retrouva son unique barrique vide, et le sol couvert de bière. Frustré, il essaya de nettoyer le désordre, mais il se souvint qu'il avait encore besoin de baratter de beurre. Il alla donc à la laiterie, et trouva assez de crème pour remplir à nouveau la baratte. Il était décidé à s'occuper de la maison mieux que sa femme.

Après avoir bien baratté pendant un moment, il se souvint que la vache à lait était encore enfermée dans la grange. Elle n'avait pas encore mangé une bouchée ni bu une goutte d'eau de l'avant-midi. L'homme pensa que le voyage était long, pour faire brouter la vache dans la prairie. Il eut une idée : au lieu d'aller dans la prairie, il pouvait faire monter la vache sur le toit de la maison. Après tout, entre les fissures du toit, l'herbe avait commencé à pousser. Ce serait un moyen de s'en débarrasser. Il posa une grande planche de bois entre le sol et le toit, et la vache monta.

De retour à l'intérieur de la maison, il se souvint que la vache aurait probablement besoin de quelque chose à boire maintenant qu'elle était en train de paître. Pour aller chercher de l'eau, il devait aller au puits. Mais, à ce moment-là, il entendit le bébé jouer dans la pièce d'à côté. La petite s'était réveillée de sa sieste ! Alors, il pensait qu'il ne pouvait sûrement pas laisser la baratte sans surveillance. Le bébé pourrait ramper dans la cuisine et la faire tomber. Il décida qu'il n'aurait plus qu'à emporter la baratte avec lui. Alors, il attacha le récipient sur son dos, prit un seau et quitta la maison. Enfin arrivé au puits, il se pencha pour remplir le seau. Mais alors, toute la crème se déversa sur ses épaules coula dans le puits.

Il était presque midi et il savait que sa femme arriverait bientôt pour le dîner. Mais il n'avait même pas encore baratté de beurre ! Il pensa que la meilleure chose à faire était de commencer à préparer une bouillie. Il remplit un pot d'eau et l'accrocha au-dessus du feu. Une pensée terrible lui traversa alors l'esprit : et si la vache tombait du toit et se cassait le cou ? Il attacha une extrémité d'une corde autour du ventre de la vache, et laissa tomber l'autre extrémité dans la cheminée. Une fois de retour dans la cuisine, il attacha le bout de la corde à sa cuisse. L'eau commençait à bouillir, mais il

devait encore moudre la farine d'avoine. Il commençait à moudre, quand la vache perdit l'équilibre et tomba du toit. En tombant, elle entraina l'homme vers le haut. Il se retrouva coincé dans la cheminée. La vache était quant à elle suspendue dans les airs, à l'extérieur de la maison, à mi-chemin entre le toit et le sol.

En arrivant du champ de foin pour diner, la femme vit la vache suspendue à la corde. Prenant aussitôt pitié d'elle, la femme coupa la corde avec sa faux. Elle entendit alors un grand cri venant de la maison…

Prénom et nom de l'élève

: ……

Questions

1) Coche la case convenable à propos de l'état de la maison à la fin de l'histoire :

À la fin de l'histoire, on comprend que :

	Vrai	Faux	On ne sait pas
a. L'homme s'est probablement blessé en tombant			
b. Le cochon s'est enfui de la maison			
c. La vache s'est cassé le cou			
d. Des flammes sortent de la cheminée			
e. La femme est fatiguée de son travail aux champs			
f. Le dîner est prêt			

2) Qu'est-ce qui est récolté dans le champ ?

a) De l'avoine
b) Du coton
c) Du foin
d) Du maïs

3) Qu'est-ce qu'il y a dans le tonneau à la cave ?

a) Du vin

b) Du cidre
c) De la bière
d) De l'eau

4) Comment la crème a-t-elle été perdue, les deux fois ?

a) L'homme a bu toute la crème et le cochon a renversé la baratte
b) Le cochon a fait tomber la baratte de crème et l'homme l'a renversée
c) Le cochon puis la vache ont renversé la crème
d) La vache s'est roulée dans la crème, puis l'homme a renversé la baratte

5) Pourquoi l'homme a attaché le bout d'une corde à sa cuisse et l'autre bout à la vache ?

a) Pour faire brouter la vache dans le pré
b) Pour faire boire la vache sur le toit
c) Pour être averti si la vache s'aventurait trop loin sur le toit
d) Pour guider la vache sur le toit

6) Coche la case de la phrase qui peut compléter le texte.

	En entrant dans la cuisine, la femme vit la vache et son mari par terre, recouverts de bouillie.
	En entrant dans la cuisine, la femme fut saisie de l'odeur forte de bouillie qui émanait de la cheminée.
	En entrant dans la cuisine, la femme vit son mari par terre, ébouillanté, et le pot roulant sur le sol.
	En entrant dans la cuisine, la femme vit le sol couvert de bière.

7) Ces phrases du texte ont été mélangées. Remets-les dans l'ordre du texte <u>en les numérotant de 1 à 5</u>. Écris les numéros dans les cases ci-dessous :

	Dans sa fureur, l'homme oublia tout ce qu'il avait commencé en bas et se mit à hurler et à poursuivre le cochon dans la maison.
	Après avoir baratté pendant un certain temps, il eut soif.
	Il posa une grande planche de bois entre le sol et le toit, et la vache monta.
	L'eau commençait à bouillir, mais il devait encore moudre la farine d'avoine.
	Mais, à ce moment-là, il entendit le bébé jouer dans la pièce d'à côté.

8) Encercle toutes les tâches que le mari <u>devait accomplir</u> cette journée-là.

- Baratter du beurre
- Récolter le foin
- Donner à manger et à boire à la vache
- Donner à manger et à boire au cochon
- S'occuper du bébé
- Faire brouter la vache dans la prairie
- Faire le dîner
- Moudre la farine d'avoine

9) Encercle les titres qui conviennent à l'histoire.

- La journée des moissonneurs
- Le mari qui devait prendre soin de la maison
- Une journée ordinaire à la ferme

- Le mari qui était très rusé
- Le barattage de la crème
- La vache et le cochon font des histoires
- Comment le cochon s'est enfui
- Un avant-midi merveilleux
- Une proposition ingénieuse
- À chacun ses talents

10) Coche la case convenable.

Pourrait-on se servir de ce texte pour :

	Oui	Non	Je ne sais pas
a. montrer comment faire de la bouillie			
b. montrer l'importance de l'empathie			
c. montrer les étapes de fabrication du beurre			
d. montrer que tenir une maison est difficile			

11) Coche la case convenable pour répondre à la question suivante. D'après toi, pourquoi ce texte a-t-il été écrit ?

	Pour montrer que les gens peuvent être très aimables
	Pour expliquer le travail des moissonneurs
	Pour décrire la saison des récoltes
	Pour montrer que la gestion de la maison est un vrai travail
	Pour démontrer l'arrogance ou l'orgueil de certaines personnes

12) D'après toi, pourquoi le mari a accepté la proposition de sa femme?

13) D'après toi, est-ce que le texte est pour l'idée que les tâches ménagères doivent être partagées dans un couple ? Pourquoi?

Merci d'avoir répondu aux questions ☺

ANNEXE H

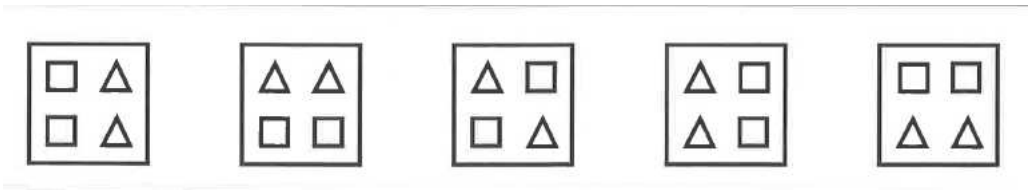
EXEMPLES D'ITEMS DES ÉPREUVES MATRICIELLES AVEC MATÉRIAU NON LINGUISTIQUE

Les items proviennent des épreuves de matrices et de classifications du CFT1-R (Weiß et Osterland, 2013)

Sous-test 5 CFT1-R : **Classifications**, 4 minutes

Cinq figures sont présentées au sujet. Il ou elle doit en sélectionner une qui est différente des autres.

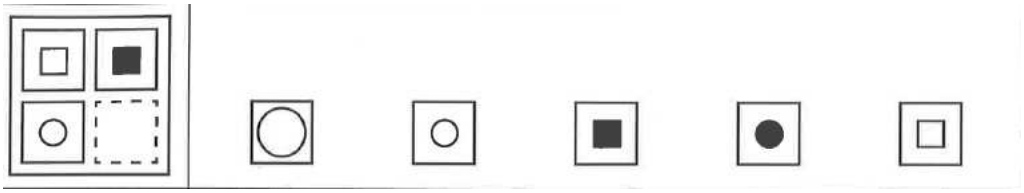
Exemple d'item :



Sous-test 6 CFT1-R : **Matrices**, 2 minutes 30 secondes

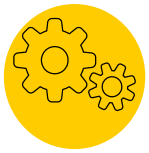
Le sujet doit compléter la matrice présentée à gauche en sélectionnant une des figures présentées à droite.

Exemple d'item :



ANNEXE I

DOCUMENT INCITATIF DESTINÉ AUX ENSEIGNANT·ES



Recherche sur le fonctionnement de la langue et la lecture

⇒ **Vous vous posez des questions sur les liens entre le fonctionnement de la langue et la lecture ?**
⇒ **Vous avez envie de connaître de nouveaux exercices sur le fonctionnement de la langue pour vos élèves ?**
⇒ **Vous voudriez collaborer avec des chercheuses de la HEP pour penser à l'enseignement de ces objets ?**

Dans cette recherche, nous nous intéressons à l'apprentissage du fonctionnement de la langue (FDL) au primaire ainsi qu'aux liens entre le raisonnement grammatical, l'oral et la lecture. Comment penser les liens entre FDL et lecture, comme recommandé par le PER ? Nous souhaitons tester de nouveaux dispositifs de type MITIC (exercices sur ordinateur) et faire progresser les élèves. Votre participation à cette recherche serait un cadeau pour nous et pour vos élèves !

Description de la recherche

Degrés : classes de 5H à 6H
Objets d'enseignement : mots de même famille, flexions grammaticales, groupes et fonctions en grammaire, compréhension en lecture
Objectifs : comprendre où se situent les difficultés des élèves dans l'analyse lexicale et grammaticale et en compréhension de textes à travers une série d'exercices*
Votre participation : rendre vos classes disponibles pour 6 séances d'exercices* faits à l'ordinateur et vous entretenir avec nous à propos de ce que vous observez chez vos élèves.

Calendrier

Incitatifs à la participation à ce projet

- Mieux comprendre les difficultés des élèves et envisager des moyens de les aider. L'élève est véritablement au centre de nos préoccupations.
- Collaborer à un projet sur la question de l'enseignement et de l'apprentissage du fonctionnement de la langue et la lecture.
- Trouver des solutions quant à vos questions sur les objets d'enseignement précités.

* Ces exercices ont été testés auprès de différentes classes au Québec, en France et en Suisse, et ont fait l'objet de travaux de recherche par la suite. Nous souhaitons que vos élèves bénéficient de ces exercices aussi ! Six séances sont nécessaires. Nous savons que votre temps en classe est précieux, et c'est pourquoi nous souhaitons collaborer étroitement avec vous. Le rythme de ces 6 séances est à votre discrétion (1 X par semaine, 2 X par semaine, ou autre configuration qui vous va mieux).

Écrivez-nous ! Nous nous ferons un plaisir de discuter du projet avec vous.
Contact : Rosalie Bourdages, rosalie.bourdages@hepl.ch, whatsapp : +15145045757
Roxane Gagnon, roxane.gagnon@hepl.ch
Site du projet : https://www.2cr2d.ch/didactique-de-loral/

hep/ haute école pédagogique vaud

ANNEXE J

FORMULAIRE DE CONSENTEMENT POUR LES PARENTS D'ÉLÈVES

Recherche sur le fonctionnement de la langue et la lecture

Nom de l'établissement scolaire
Classe, Nom de l'enseignant·e

Aux Parents des élèves de la classe
Ville, date

Chers Parents,

Je participe à une recherche sur le fonctionnement de la langue et la lecture, réalisée par des chercheures de la HEP Vaud, qui vise à mieux comprendre les difficultés des élèves en français et envisager des moyens de les aider. Les élèves de la classe auront à réaliser des petits exercices à l'ordinateur, à raison d'une fois par semaine pendant environ 5 ou 6 semaines. Les informations recueillies lors de cette rencontre seront utilisées uniquement à des fins scientifiques et de formation : la confidentialité des réponses des élèves est assurée. Les résultats de cette étude n'auront aucun impact sur les notes de votre enfant. Les avantages et les inconvénients encourus par votre enfant, s'il accepte avec votre accord de participer à notre recherche, sont comparables à ceux rencontrés lors d'une journée scolaire typique.

Pour que votre enfant participe à cette petite enquête, j'ai besoin de votre autorisation écrite, raison pour laquelle je vous adresse ces quelques lignes.

En restant à votre disposition pour tout renseignement complémentaire et en vous remerciant de votre collaboration, je vous adresse, chers Parents, mes meilleures salutations.

Nom de l'enseignant·e

Copie à :
- Direction

Autorisation de recueil de données, sous forme d'exercices sur le fonctionnement de la langue et la lecture – Formulaire de consentement

(Coupon à retourner à l'enseignant·e d'ici au 5 novembre 2019 si possible)

Nom et prénom de l'élève : .....

Je, soussigné ..... , atteste avoir pris connaissance des informations données ci-dessus et autorise que mon enfant réalise ces exercices à l'ordinateur.

OU

Je, soussigné ..... , atteste avoir pris connaissance des informations données ci-dessus, mais refuse que mon enfant réalise ces exercices à l'ordinateur.

En cas d'accord, merci de répondre encore aux questions ci-dessous :

Langue parlée à la maison .....

Langue première de la maman .....

Langue première du papa .....

Lieu et date : .....

Signature (représentant légal) : .....

ANNEXE K

STATISTIQUES DESCRIPTIVES ET CORRÉLATIONS POUR LES TEXTES DE L'ÉPREUVE AVEC TEXTES COURTS

Anova à mesures répétées

Effets intra-sujets

	Correction de la spécificité	Somme des carrés	ddl	Carrés moyens	F	p	η^2	η^2_p
Texte	Aucune	46.5	9	5.1691	73.1	<.001	0.222	0.278
	Greenhouse-Geisser	46.5	8.39	5.5435	73.1	<.001	0.222	0.278
Résidu	Aucune	120.9	1710	0.0707				
	Greenhouse-Geisser	120.9	1594.51	0.0758				

Note. Somme des carrés de type 3

Effets inter-sujets

	Somme des carrés	ddl	Carrés moyens	F	p	η^2	η^2_p
Résidu	41.8	190	0.220				

Note. Somme des carrés de type 3

Hypothèses

Tests de sphéricité

	W de Mauchly	p	ε de Greenhouse-Geisser	ε de Huynh-Feldt
Texte	0.728	0.062	0.932	0.979

Tests post hoc

Comparaisons post hoc - Texte

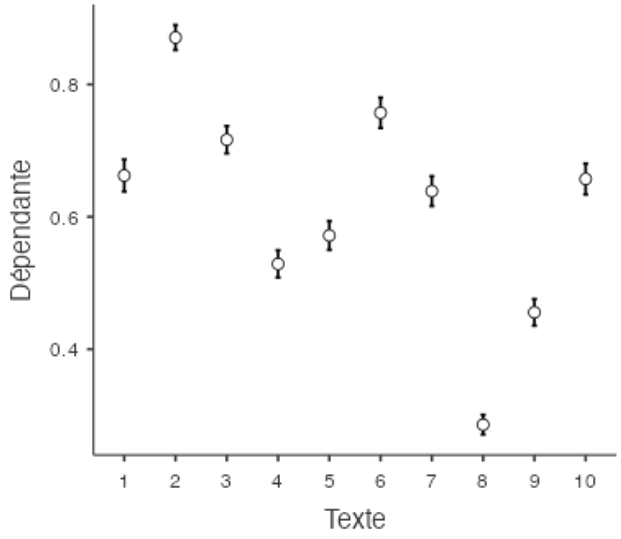
Comparaison							
Texte		**Texte**	**Différence moyenne**	**Erreur standard**	**ddl**	**t**	**p_{tukey}**
1	-	2	-0.20855	0.0261	190	-8.000	<.001
	-	3	-0.05410	0.0290	190	-1.863	0.694
	-	4	0.13351	0.0290	190	4.604	<.001
	-	5	0.09058	0.0301	190	3.012	0.084
	-	6	-0.09476	0.0298	190	-3.182	0.053
	-	7	0.02356	0.0283	190	0.833	0.998
	-	8	0.37644	0.0263	190	14.326	<.001
	-	9	0.20681	0.0267	190	7.735	<.001
	-	10	0.00524	0.0294	190	0.178	1.000
2	-	3	0.15445	0.0248	190	6.230	<.001
	-	4	0.34206	0.0246	190	13.897	<.001
	-	5	0.29913	0.0257	190	11.636	<.001
	-	6	0.11379	0.0273	190	4.173	0.002
	-	7	0.23211	0.0260	190	8.926	<.001
	-	8	0.58499	0.0242	190	24.210	<.001
	-	9	0.41536	0.0252	190	16.466	<.001
	-	10	0.21379	0.0285	190	7.510	<.001
3	-	4	0.18761	0.0244	190	7.693	<.001
	-	5	0.14468	0.0263	190	5.502	<.001
	-	6	-0.04066	0.0296	190	-1.376	0.933
	-	7	0.07766	0.0289	190	2.685	0.187
	-	8	0.43054	0.0243	190	17.690	<.001
	-	9	0.26091	0.0265	190	9.842	<.001
	-	10	0.05934	0.0278	190	2.132	0.507
4	-	5	-0.04293	0.0271	190	-1.583	0.855
	-	6	-0.22827	0.0271	190	-8.429	<.001
	-	7	-0.10995	0.0261	190	-4.214	0.002
	-	8	0.24293	0.0241	190	10.069	<.001
	-	9	0.07330	0.0264	190	2.778	0.152
	-	10	-0.12827	0.0274	190	-4.684	<.001
5	-	6	-0.18534	0.0304	190	-6.088	<.001
	-	7	-0.06702	0.0286	190	-2.346	0.365
	-	8	0.28586	0.0274	190	10.451	<.001
	-	9	0.11623	0.0285	190	4.084	0.003

Comparaisons post hoc - Texte

Comparaison							
Texte		**Texte**	**Différence moyenne**	**Erreur standard**	**ddl**	**t**	**p_{tukey}**
	-	10	-0.08534	0.0302	190	-2.830	0.134
6	-	7	0.11832	0.0262	190	4.518	<.001
	-	8	0.47120	0.0281	190	16.752	<.001
	-	9	0.30157	0.0267	190	11.310	<.001
	-	10	0.10000	0.0276	190	3.617	0.014
7	-	8	0.35288	0.0261	190	13.494	<.001
	-	9	0.18325	0.0280	190	6.544	<.001
	-	10	-0.01832	0.0292	190	-0.628	1.000
8	-	9	-0.16963	0.0253	190	-6.708	<.001
	-	10	-0.37120	0.0258	190	-14.406	<.001
9	-	10	-0.20157	0.0269	190	-7.483	<.001

Estimation des moyennes marginales

Texte



Estimation des moyennes marginales - Texte

			Intervalle de confiance à 95%	
Texte	**Moyenne**	**Erreur standard**	**Borne inf**	**Borne sup**
1	0.662	0.0242	0.615	0.710

Estimation des moyennes marginales - Texte

			Intervalle de confiance à 95%	
Texte	**Moyenne**	**Erreur standard**	**Borne inf**	**Borne sup**
2	0.871	0.0189	0.834	0.908
3	0.716	0.0207	0.676	0.757
4	0.529	0.0207	0.488	0.570
5	0.572	0.0219	0.529	0.615
6	0.757	0.0230	0.712	0.802
7	0.639	0.0225	0.594	0.683
8	0.286	0.0149	0.257	0.315
9	0.455	0.0201	0.416	0.495
10	0.657	0.0234	0.611	0.703

Matrice de corrélation pour la réussite aux différents textes de l'épreuve avec textes courts (Note. * p < .05, ** p < .01, *** p < .001)

		TXT10	TXT9	TXT8	TXT7	TXT6	TXT5	TXT4	TXT3	TXT2	TXT1
TXT10	r de Pearson	—									
	valeur p	—									
TXT9	r de Pearson	0.239 ***	—								
	valeur p	<.001	—								
TXT8	r de Pearson	0.150 *	-0.025	—							
	valeur p	0.039	0.732	—							
TXT7	r de Pearson	0.192 **	0.139	0.066	—						
	valeur p	0.008	0.055	0.363	—						
TXT6	r de Pearson	0.290 ***	0.240 ***	-0.059	0.338 ***	—					
	valeur p	<.001	<.001	0.416	<.001	—					
TXT5	r de Pearson	0.112	0.081	-0.076	0.171 *	0.079	—				
	valeur p	0.122	0.266	0.296	0.018	0.275	—				
TXT4	r de Pearson	0.231 **	0.161 *	0.107	0.272 ***	0.234 **	0.187 **	—			
	valeur p	0.001	0.026	0.140	<.001	0.001	0.010	—			
TXT3	r de Pearson	0.206 **	0.154 *	0.091	0.105	0.087	0.236 **	0.303 ***	—		
	valeur p	0.004	0.034	0.210	0.150	0.229	0.001	<.001	—		
TXT2	r de Pearson	0.105	0.163 *	-0.011	0.220 **	0.163 *	0.210 **	0.227 **	0.216 **	—	
	valeur p	0.149	0.025	0.876	0.002	0.024	0.004	0.002	0.003	—	
TXT1	r de Pearson	0.240 ***	0.283 ***	0.163 *	0.269 ***	0.206 **	0.151 *	0.173 *	0.170 *	0.288 ***	—

Matrice de corrélation pour la réussite aux différents textes de l'épreuve avec textes courts (Note. * p < .05, ** p < .01, *** p < .001)

	TXT10	TXT9	TXT8	TXT7	TXT6	TXT5	TXT4	TXT3	TXT2	TXT1
valeur p	<.001	<.001	0.024	<.001	0.004	0.036	0.017	0.018	<.001	—

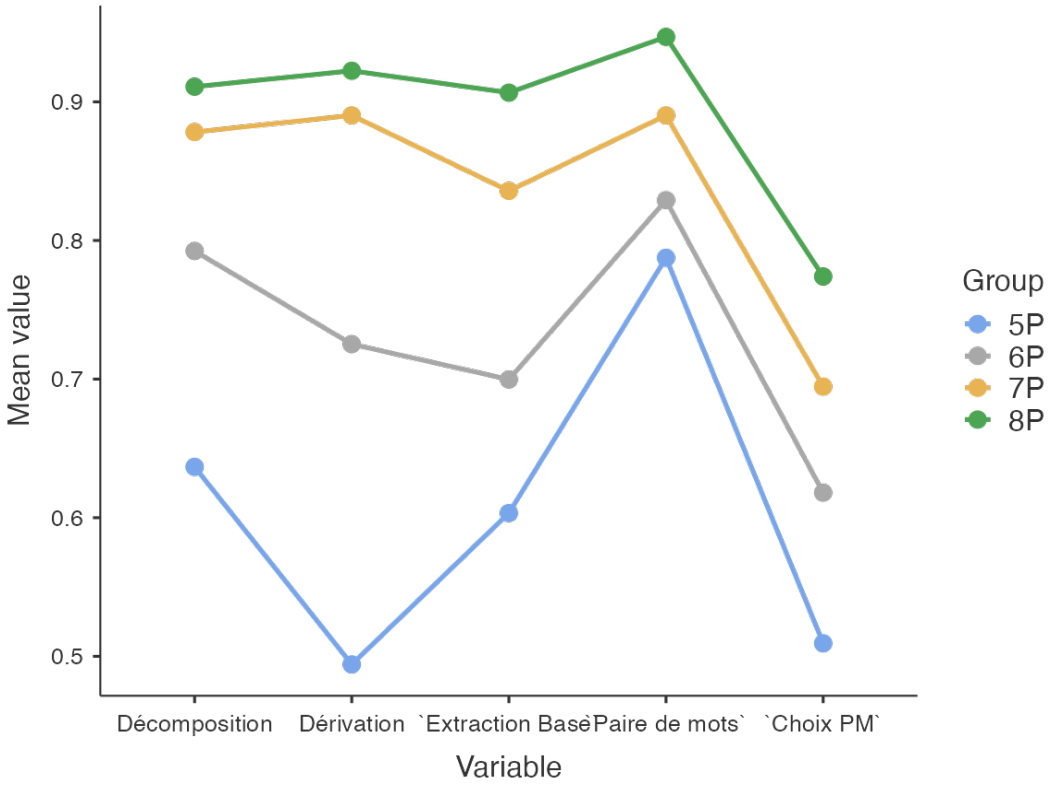
ANNEXE L

STATISTIQUES DESCRIPTIVES POUR LES ÉPREUVES MESURANT LES QUATRE GRANDS CONSTRUITS, EN FONCTION DU NIVEAU SCOLAIRE

L.1 Résultats des élèves aux épreuves de conscience morphologique, en fonction des 4 niveaux scolaires

Statistiques descriptives sur les moyennes de réussite des élèves aux tâches de conscience morphologique en fonction des 4 niveaux scolaires

	Décomposition	**Dérivation**	**Extraction Base**	**Paire de mots**	**Choix PM**
5P	0.637	0.494	0.603	0.787	0.509
6P	0.792	0.725	0.700	0.829	0.618
7P	0.878	0.890	0.836	0.890	0.695
8P	0.911	0.922	0.907	0.947	0.774

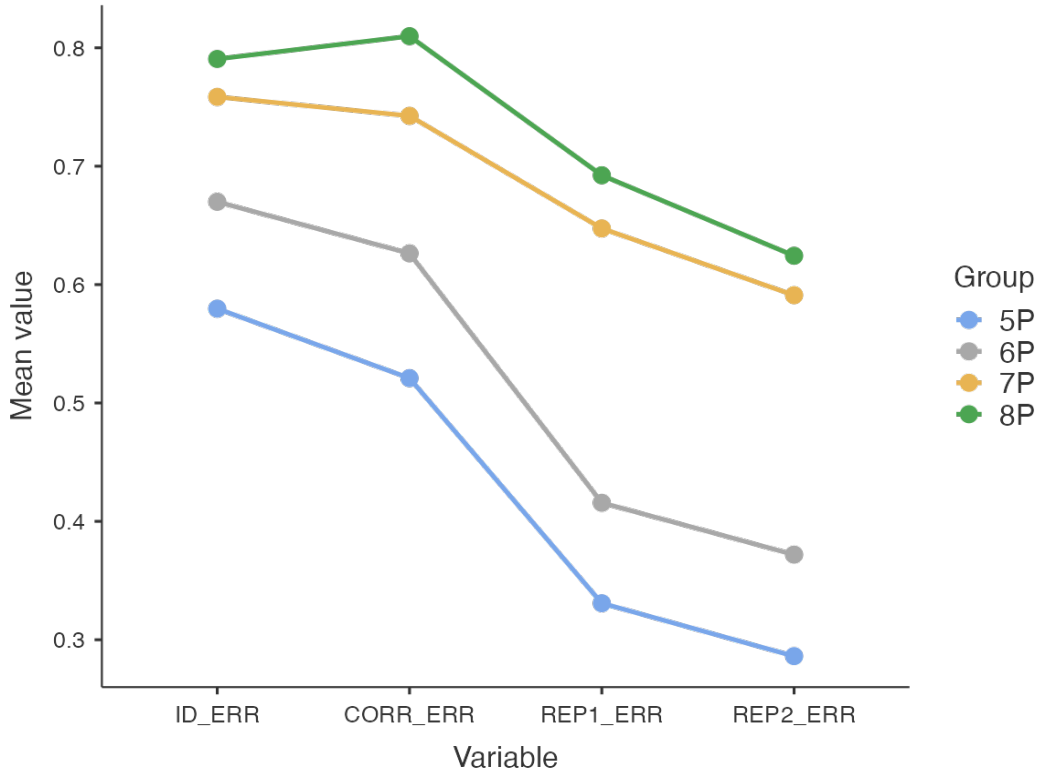


Représentation graphique des moyennes de réussite des élèves aux tâches de conscience morphologique en fonction des 4 niveaux scolaires

L.2 Résultats des élèves aux épreuves de conscience syntaxique, en fonction des 4 niveaux scolaires

Statistiques descriptives sur les moyennes de réussite des élèves aux tâches de conscience syntaxique en fonction des 4 niveaux scolaires

	Identification de l'erreur	**Correction de l'erreur**	**Réplication de l'erreur 1**	**Réplication de l'erreur 2**
5P	0.580	0.521	0.331	0.286
6P	0.670	0.626	0.416	0.372
7P	0.759	0.743	0.647	0.591
8P	0.791	0.810	0.692	0.624

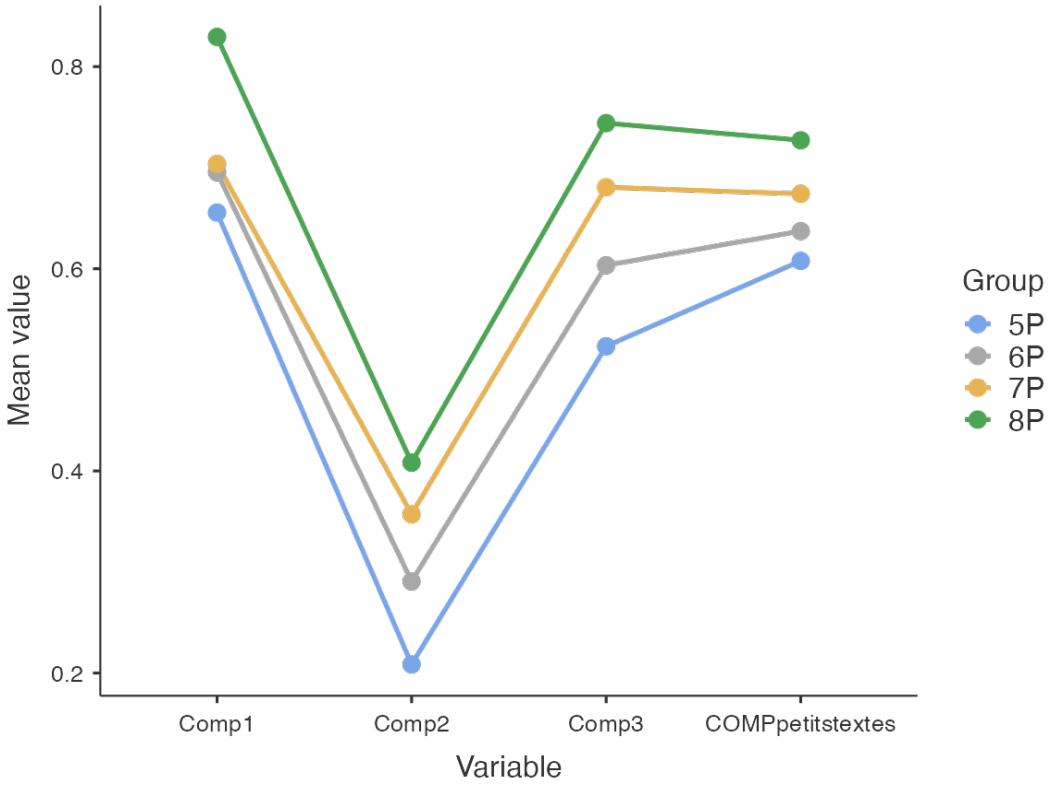


Représentation graphique des moyennes de réussite des élèves aux tâches de conscience syntaxique en fonction des 4 niveaux scolaires

L.3 Résultats des élèves aux épreuves de compréhension en lecture, en fonction des 4 niveaux scolaires

Statistiques descriptives sur les moyennes de réussite des élèves aux épreuves de compréhension en fonction des 4 niveaux scolaires

	Comp1	Comp2	Comp3	COMPpetitstextes
5P	0.656	0.209	0.523	0.608
6P	0.695	0.291	0.603	0.637
7P	0.704	0.357	0.681	0.674
8P	0.829	0.408	0.744	0.727

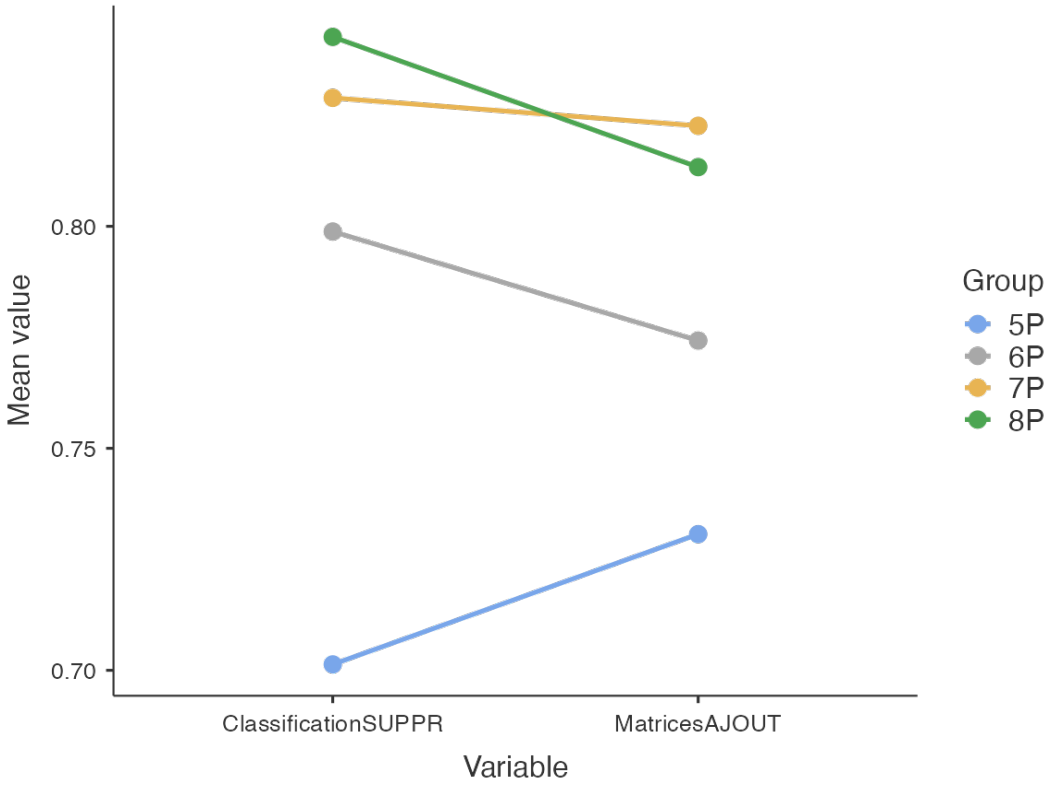


Représentation graphique des moyennes de réussite des élèves aux épreuves de compréhension en fonction des 4 niveaux scolaires

L.4 Résultats des élèves aux épreuves transformationnelles avec matériau non linguistique, en fonction des 4 niveaux scolaires

Statistiques descriptives sur les moyennes de réussite des élèves aux épreuves transformationelles avec matériau non linguistique en fonction des 4 niveaux scolaires

	ClassificationSUPPR	**MatricesAJOUT**
5P	0.701	0.731
6P	0.799	0.774
7P	0.829	0.823
8P	0.843	0.813



Représentation graphique des moyennes de réussite des élèves aux épreuves transformationelles avec matériau non linguistique en fonction des 4 niveaux scolaires

ANNEXE M

ÉPREUVES ET NOM DES VARIABLES ASSOCIÉES

Le tableau suivante présente les épreuves et les noms des variables auxquelles elles sont associées, en fonction des deux dimensions hypothétisées (unités et opérations).

Épreuves	**Unités**	**Opérations**	**Nom des variables associées**
	Compréhension en lecture		
Texte long	Compréhension littérale		Comp1
	Inférences		Comp2
	Lien entre les deux premiers niveaux		Comp3
Textes courts	Tous niveaux de compréhension confondus		COMPpetitstextes
	Conscience morphologique		
Paires de mots	Mots dérivés et pseudo-dérivés avec suffixes dérivationnels	Suppression	PaireMotsSUPP
Extraction de la base	Suffixes dérivationnels	Suppression	ExtractionBaseTotalSUPP
Dérivation	Morphèmes dérivationnels et flexionnels	Ajout	DerivationDAJOUT et DerivationFAJOUT
Décomposition	Morphèmes dérivationnels et flexionnels	Suppression	DecompositionDSUPP et DecompositionFSUPP
Choix de pseudomots suffixés D	Morphèmes dérivationnels	Ajout	PMderivAJOUT
Choix de pseudomots suffixés F	Morphèmes flexionnels	Ajout	PMflexAJOUT
	Conscience syntaxique		

Identification de l'erreur dans phrase agrammaticale M	Mot	Suppression	ID_ERREUR_PAGRAMM_M
Correction de l'erreur dans phrase agrammaticale M	Mot	Ajout	CORR_ERREUR_PAGRAMM_M
Identification phrase grammaticale, réplication 1 M	Mot	Suppression	ID_PGRAMM_REP1_M
Correction phrase grammaticale, réplication 1 M	Mot	Ajout	CORR_PGRAMM_REP1_M
Identification phrase grammaticale, réplication 2 M	Mot	Suppression	ID_PGRAMM_REP2_M
Correction phrase grammaticale, réplication 2 M	Mot	Ajout	CORR_PGRAMM_REP2_M
Identification de l'erreur dans phrase agrammaticale S	Syntagme	Suppression	ID_ERREUR_PAGRAMM_S
Correction de l'erreur dans phrase agrammaticale S	Syntagme	Ajout	CORR_ERREUR_PAGRAMM_S
Identification phrase grammaticale, réplication 1 S	Syntagme	Suppression	ID_PGRAMM_REP1_S
Correction phrase grammaticale, réplication 1 S	Syntagme	Ajout	CORR_PGRAMM_REP1_S
Identification phrase grammaticale, réplication 2 S	Syntagme	Suppression	ID_PGRAMM_REP2_S
Correction phrase grammaticale, réplication 2 S	Syntagme	Ajout	CORR_PGRAMM_REP2_S

Épreuves transformationnelles avec matériau non linguistique

1. Sous-test 6 CFT1-R : matrices	Ajout	MatricesAJOUT
2. Sous-test 5 CFT1-R : classification (intrus)	Suppression	ClassificationSUPPR

RÉFÉRENCES

Aarnoutse, C., Van Leeuwe, J., Voeten, M. et Oud, H. (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. *Reading and Writing: An Interdisciplinary Journal*, *14*, 61–89.

Akturk, A. O. et Sahin, I. (2011). Literature Review on Metacognition and its Measurement. *Procedia - Social and Behavioral Sciences*, *15*, 3731-3736. https://doi.org/10.1016/j.sbspro.2011.04.364

Alexandre-Beaumont, C., Delehaye-Boidin, C. et Le Hay, L. (s.d.). *B.L.C.R. – Batterie de Lecture et de Compréhension de Récit pour CE2, CM1 et CM2*. Ortho&logo.

Allard, D. (2008). *Connaissances lexicales et en morphologie dérivationnelle chez des apprenants adultes avancés en français langue seconde* [Mémoire de maitrise, Université du Québec à Montréal]. Montréal.

Ambridge, B. (2017). Syntactic Categories in Child Language Acquisition. Dans *Handbook of Categorization in Cognitive Science* (p. 567-580). https://doi.org/10.1016/b978-0-08-101107-2.00023-3.

Anderson, S. R. (2016). The Role of Morphology in Transformational Grammar. Dans A. Hippisley et G. Stump (dir.), *The Cambridge Handbook of Morphology* (p. 588-608). Cambridge University Press. https://doi.org/DOI:10.1017/9781139814720.021.

Anderson, S. R. (2018). A Short History of Morphological Theory. Dans J. Audring et F. Masini (dir.), *The Oxford Handbook of Morphological Theory* (p. 18-33). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199668984.013.2.

Anthony, J. L. et Lonigan, C. J. (2004). The Nature of Phonological Awareness: Converging Evidence From Four Studies of Preschool and Early Grade School Children. *Journal of Educational Psychology*, *96*(1), 43-55. https://doi.org/10.1037/0022-0663.96.1.43

Apel, K. (2014). A Comprehensive Definition of Morphological Awareness. *Topics in Language Disorders*, *34*(3), 197-209. https://doi.org/10.1097/tld.0000000000000019

Apel, K., Diehm, E. et Apel, L. (2013). Using Multiple Measures of Morphological Awareness to Assess its Relation to Reading. *Topics in Language Disorders*, *33*(1), 42-56. https://doi.org/10.1097/TLD.0b013e318280f57b

Apel, K., Petscher, Y. et Henbest, V. S. (2021). *Morphological Awareness Test for Reading and Spelling (MATRS): Technical Report*.

Apothéloz, D. (2002). *La construction du lexique français. Principes de morphologie dérivationnelle*. Ophrys. https://halshs.archives-ouvertes.fr/halshs-00005964

Armand, F. (2000). Le rôle des capacités métalinguistiques et de la compétence langagière orale dans l'apprentissage de la lecture en français langue première et seconde. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, *56*, 469-495.

Armand, F. (2005). Capacités métalinguistiques d'élèves immigrants nouvellement arrivés en situation de grand retard scolaire. *Revue des sciences de l'éducation*, *31*(2), 441-469.

Arnauld, A. et Lancelot, C. (1660/1810). *Grammaire générale et raisonnée de Port-Royal*. Bossange et Masson.

Arseneau, R. (2016). *Coopérer pour apprendre la syntaxe écrite : recherche-action sur la phrase subordonnée relative en troisième secondaire* [Thèse de doctorat, Université de Montréal]. Montréal.

Arseneau, R., Foucambert, D. et Lefrançois, P. (2016). Quand les arbres font obstacle… décloisonnement des espaces et apaisement de tensions dans une recherche-action en didactique de la grammaire. Dans *Questionner l'espace* (p. 149-162). https://doi.org/10.4000/books.septentrion.19770.

Astolfi, J.-P., Darot, É., Ginsburger-Vogel, Y. et Toussaint, J. (2008). Chapitre 14. Problème et situation-problème. Dans *Mots-clés de la didactique des sciences* (p. 137-146). De Boeck Supérieur. https://doi.org/10.3917/dbu.astol.2008.01.0137.

Atkinson, R. et Shiffrin, R. (1968). Human memory: A proposed system and its control processes. Dans K. Spence et J. Spence (dir.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (vol. 2, p. 89-195). Academic Press.

Audion, L. (2016). Pour un apprentissage de la réflexion métalinguistique à l'école élémentaire. *Pratiques*(169-170). https://doi.org/10.4000/pratiques.3050

Baddeley, A. (1986). *Working memory*. Oxford University Press.

Baker, L. (1989). Metacognition, Comprehension Monitoring, and the Adult Reader. *Educational Psychology Review*, *1*(1), 3-38.

Baker, L. et Brown, A. L. (1984). Metacognitive skills and reading. Dans P. D. Pearson, R. Barr, K. M. L. et P. B. Mosenthal (dir.), *Handbook of Reading Research* (p. 353-394). Longman.

Barac, R., Bialystok, E., Castro, D. C. et Sanchez, M. (2014). The Cognitive Development of Young Dual Language Learners: A Critical Review. *Early Child Res Q*, *29*(4), 699-714. https://doi.org/10.1016/j.ecresq.2014.02.003

Batterink, L. et Neville, H. J. (2013). The human brain processes syntax in the absence of conscious awareness. *The Journal of Neuroscience*, *33*(19), 8528-8533. https://doi.org/10.1523/JNEUROSCI.0618-13.2013

Baudrit, A. (2007). Chapitre 1. Les bases scientifiques de l'apprentissage collaboratif. Dans *L'apprentissage collaboratif* (p. 13-34). De Boeck Supérieur.

Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, *25*, 464-469.

Benveniste, E. (1974). *Problèmes de linguistique générale II*. Gallimard.

Berardi-Coletta, B., Buyer, L. S., Dominowski, R. L. et Rellinger, E. R. (1995). Metacognition and problem solving: A process-oriented approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 205-223. https://doi.org/10.1037/0278-7393.21.1.205

Berger, J.-L. (2021). Analyse factorielle exploratoire et analyse en composantes principales: guide pratique. *hal-03436771v2*.

Berger, M. J. et Desrochers, A. (2011). *L'évaluation de la littératie*. Presses de l'universté d'Ottawa, Canada.

Berko, J. (1958). The Child's Learning of English Morphology. *Word*, *14*, 150-177.

Bernstein, S. E., Flipse, J. L., Jin, Y. et Odegard, T. N. (2020). Word and sentence level tests of morphological awareness in reading. *Reading and Writing*. https://doi.org/10.1007/s11145-020-10024-6

Berthiaume, R., Besse, A.-S. et Daigle, D. (2010). L'évaluation de la conscience morphologique: proposition d'une typologie des tâches. *Language Awareness*, *19*(3), 153-170. https://doi.org/10.1080/09658416.2010.482992

Besse, A.-S., Marec-Breton, N. et Demont, E. (2010). Développement métalinguistique et apprentissage de la lecture chez les enfants bilingues. *Enfance*, *2010*(02), 167. https://doi.org/10.4074/s001375451000203x

Besse, H. (2019). Quelques réflexions sur l'histoire de ce qu'on dénomme exercices en didactique des langues. *Documents pour l'histoire du français langue étrangère ou seconde*(62-63), 43-67. https://doi.org/10.4000/dhfles.6169

Bestgen, Y. (2007). Représentation de l'espace et du temps dans le modèle situationnel construit par un lecteur. *Discours, Revue de linguistique, psycholinguistique et informatique*(1). https://doi.org/10.4000/discours.100

Bialystok, E. (1978). A Theoretical Model of second Language Learning. *Language Learning*, *28*(1), 69-83.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child development*, *57*, 498-510.

Bialystok, E. (1988a). Aspects of linguistic awareness in reading comprehension. *Applied Psycholinguistics*, *9*, 123-139. https://doi.org/10.1017/S0142716400006779

Bialystok, E. (1988b). Levels of Bilingualism and Levels of Linguistic Awareness. *Developmental Psychology*, *24*(4), 560-567.

Bialystok, E. (1993). Metalinguistic awareness: The development of children's representations of language. Dans C. P. A. F. Garton (dir.), *Systems of representation in children: Development and use* (p. 211-233). John Wiley & Sons.

Bialystok, E. (2001a). *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge University Press.

Bialystok, E. (2001b). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, *21*, 169-181.

Bialystok, E., Majumder, S. et Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, *24*, 27-44.

Bialystok, E., Peets, K. F. et Moreno, S. (2014). Producing bilinguals through immersion education: Development of metalinguistic awareness. *Applied*

Psycholinguistics, *35*(1), 177-191. https://doi.org/10.1017/S0142716412000288

Bialystok, E. et Ryan, E. B. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly*, *31*(3), 229-251.

Bianco, M., Pellenq, C., Lambert, E., Bressoux, P., Lima, L. et Doyen, A.-L. (2012). Impact of early code-skill and oral-comprehension training on reading achievement in first grade. *Journal of Research in Reading*, *35*(4), 427-455. https://doi.org/10.1111/j.1467-9817.2010.01479.x

Bigras, I. (2019). *Les habiletés métasyntaxiques et les habiletés d'organisation des connaissances : y a-t-il des processus cognitifs communs?* [Mémoire de maitrise, Université du Québec à Montréal]. Montréal.

Blakmore, A. M. et Pratt, C. (1997). Grammatical Awareness and Reading in Grade 1 Children. *Merrill-Palmer Quarterly*, *43*(4), 567-590.

Blanc, A.-C., Gagnon, R. et Colognesi, S. (2023). L'articulation oral-écrit dans les dispositifs d'enseignement de la production de récits au primaire. *Le français aujourd'hui*, *222*, 53-67.

Bloomfield, L. (1933). *Language*. Holt, Rinehart & Winston.

Boivin, M.-C. (2009). Jugements de grammaticalité et manipulations syntaxiques dans le travail en classe d'élève du secondaire. Dans J. Dolz et C. Simard (dir.), *Pratiques d'enseignement grammatical. Points de vue de l'enseignant et de l'élève* (p. 179-208). Presses de l'Université Laval.

Boivin, M.-C. (2014). Quand les élèves « font de la grammaire » en classe : analyse d'interventions métalinguistiques d'élèves du secondaire. *Repères*, *49*, 131-145.

Borel-Maisonny, S. (1956). Méthode de lecture - Atlas. *Bulletin mensuel de la Société Alfred Binet*, *427*, 41-76.

Bouchard, D. (2019). La linguistique en toute simplicité / Linguistics, simply. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *64*(02), 360-405. https://doi.org/10.1017/cnj.2018.35

Bourdages, R. (2014). *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 de maitrise, Université du Québec à Montréal].

Bourdages, R. et Foucambert, D. (2018). Regard sur les dimensions sous-jacentes dans les épreuves de conscience morphologique. *Congrès Mondial de Linguistique Française, SHS Web of Conferences*, *46*. https://doi.org/10.1051/shsconf/20184610012

Bourdages, R., Gagnon, R., Capt, V. et Foucambert, D. (2021). De la tâche à l'activité, de la mesure à l'enseignement : considérations méthodologiques sur les habiletés métasyntaxiques en psycholinguistique et en didactique de la grammaire. *Education & Didactique*, *15*(2), 49-76.

Bowey, J. A. (1986a). Syntactic awareness and verbal performance from preschool to fifth grade. *Journal of Psycholinguistic Research*, *15*(4), 285-308. https://doi.org/10.1007/BF01067676

Bowey, J. A. (1986b). Syntactic awareness in relation to reading skill and ongoing reading comprehension monitoring. *Journal of Experimental Child Psychology*, *41*(2), 282-299. https://doi.org/10.1016/0022-0965(86)90041-X

Bowey, J. A. (1994). Grammatical awareness and learning to read: A critique. Dans *Literacy acquisition and social context* (p. 122-149). Harvester Wheatsheaf/Prentice Hall.

Bowey, J. A. (2005). Predicting individual differences in learning how to read. Dans M. J. Snowling et C. Hulme (dir.), *The Science of Reading: A Handbook* (p. 55-172). Blackwell Publishing.

Bowey, J. A. et Patel, R. K. (1988). Metalinguistic ability and early reading achievement. *Applied Psycholinguistics*, *9*, 367-383.

Braine, M. D. S. (1992). What sort of innate structure is needed to "bootstrap" into syntax? *Cognition*, *45*, 77-100.

Bratlie, S. S., Brinchmann, E. I., Melby-Lervåg, M. et von Koss Torkildsen, J. (2022). Morphology—A Gateway to Advanced Language: Meta-Analysis of Morphological Knowledge in Language-Minority Children. *Review of Educational Research*, *92*(4), 614-650. https://doi.org/10.3102/00346543211073186

Bressoux, P. (2017). Practice-based research : une aporie et des espoirs. *Éducation et didactique*, *11*(3), 123-134. https://doi.org/10.4000/educationdidactique.2870

Brimo, D., Apel, K. et Fountain, T. (2015). Examining the contributions of syntactic awareness and syntactic knowledge to reading comprehension. *Journal of Research in Reading*, 1-18. https://doi.org/10.1111/1467-9817.12050

Brimo, D., Lund, E. et Sapp, A. (2018). Syntax and reading comprehension: a meta-analysis of different spoken-syntax assessments. *International Journal of Language & Communication Disorders*, *53*(3), 431-445. https://doi.org/10.1111/1460-6984.12362

Bronckart, J.-P. (1996). Genres de textes, types de discours et opérations psycholinguistiques. *Voies Livres, Pratiques et apprentissages de l'écrit*(78), 1-20.

Bronckart, J.-P. (2004). *Didactique de la grammaire*. DIP. https://archive-ouverte.unige.ch/unige:37753

Bronckart, J.-P. (2019). *Théorie du langage : une introduction critique*. Mardaga.

Bronckart, J.-P. et Chiss, J.-L. (1990). Linguistique, psycholinguistique et didactique du "FLM". Réflexions à partir d'un cursus de formation d'enseignants. *Repères*, *1*(1), 19-44. https://doi.org/10.3406/reper.1990.1991

Brown, A. L. (1987). Metacognition, Executive control, Self-Regulation and Other More Mysterious Mechanisms. Dans F. E. Weinert et R. H. Kluwe (dir.), *Metacognition, Motivation and Understanding*. Lawrence Erlbaum Associates.

Brown, J. I., Fishco, V. V. et Hanna, G. (1993). *Nelson-Denny Reading Test*. Riverside.

Bruner, J. S. (1973). Organization of Early Skilled Action. *Child Development*, *44*(1), 1-11.

Bruner, J. S. (1983). *Child's Talk. Learning to Use Language*. Oxford University Press.

Bruner, J. S. (2000). Chapitre XII. Piaget et Vygotsky. Célébrons la divergence. Dans *L'esprit piagétien*. https://doi.org/10.3917/puf.houde.2000.01.0237.

Bulea Bronckart, E. (2015). *Didactique de la grammaire : une introduction illustrée*. Université de Genève.

Bulea Bronckart, E. (2022). De l'action sur la langue à la langue en action : une réflexion sur et à partir des manipulations syntaxiques. *Forum Lecture*, *1*.

Bulea Bronckart, E. et Gagnon, R. (2017). *Former à l'enseignement de la grammaire*. Presses Universitaires du Septentrion.

Bulea Bronckart, E. et Garcia-Debanc, C. (2021). *L'étude du fonctionnement de la langue dans la discipline Français : quelles articulations ?* Presses universitaires de Namur.

Bulea Bronckart, E., Marmy Cusin, V. et Panchout-Dubois, M. (2017). Les exercices grammaticaux dans le cadre de l'enseignement rénové du français : usages, problèmes, perspectives. *Repères*(56), 131-149.

Butterworth, B. (1983). *Language Production, vol. II: Development, Writing and Other Language Processes*. Academic Press.

Byrd, J. K. (2008). *A Monte Carlo Investigation of Three Different Estimation Methods in Multilevel Structural Equation Modeling Under Conditions of Data Nonnormality and Varied Sample Sizes* [Thèse de doctorat, Texas A&M University].

Byrne, B. M. (2009). *Structural Equation Modeling With AMOS : Basic Concepts, Applications, and Programming* (2 éd.). Routledge.

Cain, K. (2007). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics*, *28*(04), 679-694.

Cain, K. et Oakhill, J. (dir.). (2007). *Children's comprehension problems in oral and written language: A cognitive perspective*. Guilford Press.

Cain, K., Oakhill, J. et Bryant, P. (2004). Children's Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills. *Journal of Educational Psychology*, *96*(1), 31-42.

Campion, N. et Rossi, J.-P. (1999). Inférences et compréhension de texte. *L'année psychologique*, *99*(3), 493-527.

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. Dans L. B. Feldman (dir.), *Morphological aspects of language processing* (p. 189–209). Erlbaum.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, *12*, 169–190.

Carlisle, J. F. et Nomanbhoy, D. M. (1993). Phonological and morphological awareness in first graders. *Applied Psycholinguistics*, *14*, 177-195.

Carlisle, J. F. et Stone, C. A. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly*, *40*(4), 428-449. https://doi.org/10.1598/rrq.40.4.3

Carpenter, P. A., Just, M. A. et Schell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven's progressive matrices test. *Psychological Review*, *97*, 404–431.

Carroll, J. M., Snowling, M. J., Stevenson, J. et Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, *39*(5), 913-923. https://doi.org/10.1037/0012-1649.39.5.913

Casalis, S., Colé, P. et Sopo, D. (2004). Morphological awareness in developmental dyslexia. *Annals of Dyslexia*, *54*(1), 114-138. https://doi.org/10.1007/s11881-004-0006-z

Casalis, S., Dusautoir, M., Colé, P. et Ducrot, S. (2009). Morphological effects in children word reading: A priming study in fourth graders. *British Journal of Developmental Psychology*, *27*(3), 761-766. https://doi.org/10.1348/026151008x389575

Casalis, S. et Louis-Alexandre, M.-F. (2000). Morphological analysis, phonological analysis and learning to read French: a longitudinal study. *Reading and Writing: An Interdisciplinary Journal*, *12*, 303–335.

Casalis, S., Mathiot, E., Bécavin, A.-S. et Colé, P. (2002). Conscience morphologique chez des apprentis lecteurs tout-venant et en difficultés. (p. 57-66). Actes du 3e Forum de morphologie de Villeneuve d'Ascq, 19-21 septembre.

Cattell, R. B. (1959/1970). *Epreuve d'intelligence sans apport culturel (Mesure du facteur « G »), Echelle III, Variantes A et B*. Institut de Recherches psychologiques, inc.

Cattell, R. B., Krug, S. E. et Barton, K. (1973a). *Measuring Intelligence with the Culture Fair Tests, Manual for Scales 2 and 3*. IPAT.

Cattell, R. B., Krug, S. E. et Barton, K. (1973b). *Technical Supplement for the Culture Fair Intelligence Tests, Scales 2 and 3*. IPAT.

Chartrand, S.-G. (2012a). *Les manipulations syntaxiques : de précieux outils pour comprendre le fonctionnement de la langue et corriger un texte*. Centre collégial de développement de matériel didactique du Collège de Maisonneuve.

Chartrand, S.-G. (2012b). Quelles finalités pour l'enseignement grammatical à l'école? Une analyse des points de vue des didacticiens du français depuis 25 ans. *Formation et profession*, *20*(3), 48-59. https://doi.org/10.18162/fp.2012.222

Chartrand, S.-G. (2013). *Les manipulations syntaxiques : de précieux outils pour étudier la langue et corriger ses textes*. CCMDD.

Chartrand, S.-G. et De Pietro, J.-F. (2012). Pour une harmonisation des terminologies grammaticales scolaires de la francophonie : quels critères pour quelles finalités ? *Enjeux*(84), 5-31.

Cheung, G. W. et Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233-255. https://doi.org/10.1207/s15328007sem0902_5

Chevallard, Y. (1985). *La Transposition didactique : du savoir savant au savoir enseigné*. La pensée sauvage.

Chi, M. T. et Glaser, R. (1985). Problem-Solving Ability. Dans R. Sternberg (dir.), *Human Abilities: An information processing approach* (p. 227-248). Freeman.

Chiape, P. et Siegel, L. S. (1999). Phonological awareness and reading acquisition in English and Punjabi-speaking Canadian children. *Journal of Educational Psychology*, *91*, 20-28.

Chiss, J.-L. et David, J. (2011). L'étude de la langue dans la formation des enseignants de français. *Le français aujourd'hui*, *5*(HS01). https://doi.org/10.3917/lfa.hs01.0093

Chiss, J.-L. et David, J. (2018). *Didactique du français. Enjeux disciplinaires et étude de la langue*. Armand Colin. https://doi.org/10.3917/arco.chiss.2018.01

Chomsky, N. (1955/1975). *The logical structure of linguistic theory*. Springer.

Chomsky, N. (1957). *Syntactic structures*. Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.

Chomsky, N. et Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

CIIP. (2006). *Rapport annuel de la Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin*. Neuchâtel.

CIIP. (2010-2016). *Plan d'études romand*. Neuchâtel, Suisse : Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin.

Cobb, T. (2013). *The Compleat Lexical Tutor, v.4.* lextutor.ca

Cobb, T. (2019). *Compleat Lexical Tutor Web VP v.2.5.* https://www.lextutor.ca/vp/comp/

Cohen, C. (2011). *Input Factors, Language Experiences and Metalinguistic Awareness in Bilingual Children* [Thèse de doctorat, University of Salford]. Salford, UK.

Colé, P., Royer, C., Leuwers, C. et Casalis, S. (2004). Les connaissances morphologiques dérivationnelles et l'apprentissage de la lecture chez l'apprenti-lecteur français du CP au CE2. *L'année psychologique*, *104*(4), 701-750.

Coltheart, M. (2005). Modeling Reading: The Dual-Route Approach. Dans M. J. Snowling et C. Hulme (dir.), *The Science of Reading: A Handbook* (p. 6–23). https://doi.org/10.1002/9780470757642.ch1.

Coltheart, M., Rastle, K., Perry, C., Langdon, R. et Ziegler, J. (2001). DRC: A Dual Rote Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review*, *108*(1), 204-256.

Compton, D. L., Miller, A. C., Elleman, A. M. et Steacy, L. M. (2013). Have We Forsaken Reading Theory in the Name of "Quick Fix" Interventions for Children With Reading Disability? *Scientific Studies of Reading*, *18*(1), 55-73. https://doi.org/10.1080/10888438.2013.836200

CTB/McGraw-Hill. (2008). *Test of Adult Basic Education (TABE)*. CTB/McGraw-Hill.

Culioli, A. (1968). La formalisation en linguistique. *Cahiers pour l'analyse*, *9*(16), 106-117.

Dai, S. (2021). Handling Missing Responses in Psychometrics: Methods and Software. *Psych*, *3*(4), 673-693. https://doi.org/10.3390/psych3040043

Daunay, B. (2007). Le sujet lecteur : une question pour la didactique du français. *Le français aujourd'hui*, *157*, 43-51. https://doi.org/https://doi.org/10.3917/lfa.157.0043

De Clercq, B. et Housen, A. (2019). The development of morphological complexity: A cross-linguistic study of L2 French and English. *Second Language Research*, *35*(1), 71-97. https://doi.org/10.1177/0267658316674506

De Pietro, J.-F. (2009). Pratiques métalangagières et émergence d'une posture « grammaticale ». Dans J. Dolz et C. Simard (dir.), *Pratiques d'enseignement grammatical. Points de vue de l'enseignant et de l'élève* (p. 15-47). Presses de l'Université Laval.

de Villiers, P. A. et de Villiers, J. G. (1972). Early judgments of semantic and syntactic acceptability by children. *Journal of Psycholinguistic Research*, *1*(4), 299-310. https://doi.org/10.1007/BF01067785

de Weck, G. et Marro, P. (2010). *Les troubles du langage chez l'enfant: description et évaluation*. Masson.

Deacon, S. H. et Kieffer, M. (2018). Understanding how syntactic awareness contributes to reading comprehension: Evidence from mediation and longitudinal models. *Journal of Educational Psychology*, *110*(1), 72-86. https://doi.org/10.1037/edu0000198

Deacon, S. H., Kieffer, M. J. et Laroche, A. (2014). The Relation Between Morphological Awareness and Reading Comprehension: Evidence From Mediation and Longitudinal Models. *Scientific Studies of Reading*, *18*(6), 432-451.

Deacon, S. H. et Kirby, J. R. (2004). Morphological awareness: Just "more phonological"? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, *25*, 223-238.

Deacon, S. H., Parrila, R. et Kirby, J. R. (2006). Processing of derived forms in high-functioning dyslexics. *Annals of Dyslexia*, *56*(1), 103-128. https://doi.org/10.1007/s11881-006-0005-3

Deacon, S. H., Parrila, R. et Kirby, J. R. (2008). A review of the evidence on morphological processing in dyslexics and poor readers: A strength of weakness? Dans *Learning, Cognition and Dyslexia* (p. 192-211). https://doi.org/10.4135/9780857020987.n10.

Demetriou, A., Makris, N., Spanoudis, G., Kazi, S., Shayer, M. et Kazali, E. (2018). Mapping the Dimensions of General Intelligence: An Integrated Differential-Developmental Theory. *Human Development*, *61*(1), 4-42. https://doi.org/10.1159/000484450

Demont, E., Gaux, C. et Gombert, J. É. (2006). Bilan métalinguistique. Dans P. Masson (dir.), *Les bilans de langage et de voix*.

Demont, E. et Gombert, J. É. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology*, *66*, 315-332.

Depallens, V. (2020). *Empathie, éthique et tension narrative. Modélisation des émotions en jeu dans la lecture littéraire scolaire* [Thèse de doctorat, Université de Lausanne]. Lausanne.

Desrochers, A., Kirby, J. R., Thompson, G. L. et Fréchette, S. (2009). Le rôle de la conscience phonologique dans l'apprentissage de la lecture. *Revue du Nouvel-Ontario*(34), 59.

Diaz, R. M. (1985). Bilingual Cognitive Development: Addressing Three Gaps in Current Research. *Child Development*, *56*(6), 1376-1388.

Dick, F., Gernsbacher, M. A. et Roberson, R. R. (2002). Syntactic processing in high- and low-skill comprehenders working under normal and stressful conditions. *CRL Newsletter*, *14*(1), 3-17.

Diependaele, K., Sandra, D. et Grainger, J. (2009). Semantic transparency and masked morphological priming: the case of prefixed words. *Memory & Cognition*, *37*(6), 895-908. https://doi.org/10.3758/MC.37.6.895

Dolz, J. et Gagnon, R. (2008). Le genre du texte, un outil didactique pour développer le langage oral et écrit. *Pratiques*(137-138), 179-198. https://doi.org/10.4000/pratiques.1159

Dolz, J. et Simard, C. (2009). Introduction. Dans J. Dolz et C. Simard (dir.), *Pratiques d'enseignement grammatical. Points de vue de l'enseignant et de l'élève* (p. 1-14). Presses de l'Université Laval.

Dolz, J. et Toulou, S. (2008). De la macrostructure de la séquence d'enseignement du texte d'opinion à l'analyse des interactions didactiques. *Travail et formation en éducation*, *1*. http://journals.openedition.org/tfe/596

Doudin, P.-A. et Martin, D. (1999). Vers une psychologie métacognitive. Dans P.-A. Doudin, D. Martin et O. Albanese (dir.), *Métacognition et éducation* (p. 3-29). Peter Lang.

Duncan, L. G., Casalis, S. et Colé, P. (2009). Early metalinguistic awareness of derivational morphology: Observations from a comparison of English and French. *Applied Psycholinguistics*, *30*(03), 405-440. https://doi.org/10.1017/s0142716409090213

Ehri, L. C. (2005). Development of Sight Word Reading: Phases and Findings. Dans M. J. Snowling et C. Hulme (dir.), *The Science of Reading: A Handbook* (p. 135–154). Blackwell.

Ehrlich, M.-F. et Tardieu, H. (1993). *Modèles mentaux, modèles de situation et compréhension de textes*. Masson.

El-Bialy, R., Gagné, C. L. et Spalding, T. L. (2013). Processing of English compounds is sensitive to the constituents' semantic transparency. *The Mental Lexicon*, *8*, 75-95. https://doi.org/DOI:io.iO75/ml.8.i.O4elb

Elalouf, M.-L. et Péret, C. (2009). Pratiques d'observation de la langue en France. Quelles évolutions? Quels obstacles? Dans J. Dolz et C. Simard (dir.), *Pratiques d'enseignement grammatical, point de vue de l'enseignant et de l'élève* (p. 49-72). Presses universitaires de Laval.

Elalouf, M.-L., Péret, C. et Gourdet, P. (2017). Les pratiques d'enseignement de la grammaire en France. Quels liens entre formation, utilisation des manuels et pratiques effectives ? Dans E. Bulea Bronckart et R. Gagnon (dir.), *Former à l'enseignement de la grammaire* (p. 45-67). Septentrion.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*(8), 709–724. https://doi.org/https://doi.org/10.1037/0003-066X.49.8.709

Evans, J. S. et Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223-241. https://doi.org/10.1177/1745691612460685

Fajgelj, S., Bala, G. et Kati, R. (2010). Latent Structure of Raven's Colored Progressive Matrices. *Collegium anthropologicum*, *34*(3), 1015–1026.

Farnia, F. et Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, *36*(4), 389-421. https://doi.org/10.1111/jrir.12003

Farrall, M. L. (2012). Chapter 12 : Comprehension. Dans *Reading Assessment* (p. 233-258). John Wiley & Sons, Inc.

Fayol, M. (2003). La compréhension : évaluation, difficultés et interventions. Conférence de consensus sur l'enseignement de la lecture à l'école primaire, Paris.

Fayol, M., Gombert, J. E., Lecocq, P., Sprenger-Charolles, L. et Zagar, D. (1992). *Psychologie cognitive de la lecture*. Presses universitaires de France.

Fernandez-Garcia, M. P., Vallejo-Seco, G., Livacic-Rojas, P. et Tuero-Herrero, E. (2018). The (Ir)Responsibility of (Under)Estimating Missing Data. *Frontiers in Psychology*, *9*, 556. https://doi.org/10.3389/fpsyg.2018.00556

Ferrand, L. et Spinelli, E. (2005). *Psychologie du langage*. Armand Colin.

Field, A., Miles, J. et Field, Z. (2012). *Discovering Statistics Using R*. SAGE.

Fisher, C. et Nadeau, M. (2014). Usage du métalangage et des manipulations syntaxiques au cours de dictées innovantes dans des classes du primaire. *Repères*(49), 169-191. https://doi.org/10.4000/reperes.742

Flavell, J. H. (1976). Metacognitive Aspects of Problem Solving. Dans L. B. Resnick (dir.), *The Nature of Intelligence* (p. 231-235). Earlbaum.

Flavell, J. H. (1978). Metacognitive Development. Dans J. M. Scandura et C. J. Brainerd (dir.), *Structural/Process Theories of Complex Human Behavior* (p. 213-245). Sijthoff and Noordhoff.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/https://doi.org/10.1037/0003-066X.34.10.906

Flavell, J. H. (1985). *Cognitive development*. Prentice-Hall.

Flavell, J. H. (1987). Speculations About dhe Nature and Development of Metacognition. Dans F. E. Weinert et R. H. Kluwe (dir.), *Metacognition, Motivation and Understanding*. Lawrence Erlbaum Associates.

Fobre, M. (1997). Pensée pédagogique et modèles philosophiques : le cas de la situation-problème. *Revue française de pédagogie*, 49-58.

Fodor, J. A. (1981). *Representations: Philosophical essays on the foundations of cognitive science*. The Harvester Press Limited.

Fodor, J. A. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. MIT Press.

Follmer, D. J. (2017). Executive Function and Reading Comprehension: A Meta-Analytic Review. *Educational Psychologist*, *53*(1), 42-60. https://doi.org/10.1080/00461520.2017.1309295

Foorman, B. R., Koon, S., Petscher, Y., Mitchell, A. et Truckenmiller, A. (2015). Examining General and Specific Factors in the Dimensionality of Oral Language and Reading in 4th-10th Grades. *Journal of Educational Psychology*, *107*(3), 884-899. https://doi.org/10.1037/edu0000026

Foorman, B. R., Petscher, Y. et Herrera, S. (2018). Unique and common effects of decoding and language factors in predicting reading comprehension in grades 1–10. *Learning and Individual Differences*, *63*, 12-23. https://doi.org/10.1016/j.lindif.2018.02.011

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 de doctorat, Université du Québec à Montréal]. Montréal.

Fortier, V., Bell, P., Perron, V. et Gauvin, I. (2019). Réflexions métalinguistiques et correction d'erreurs en anglais langue seconde : étude des types de réflexions et relation avec l'exactitude des corrections. *Canadian Modern Language Review*, *75*(1), 3-22.

Foucambert, D. (2003). *Syntaxe, vision parafovéale et processus de lecture : contribution du modèle structural à la pédagogie* [Thèse, Université Grenoble II]. Grenoble.

Foucambert, D. (2008). L'habileté syntaxique est-elle un gage d'une bonne compréhension en lecture? Analyse des performances de lecteurs de 8-9 ans. Congrès Mondial de Linguistique Française - CMLF 2008, Paris. Institut de Linguistique Française.

Foucambert, D. et Bourdages, R. (2017). *Logix©*. (version 1.0).

Foucambert, D. et Redmond, L. (2018). *Entrainer la syntaxe pour améliorer la compréhension en lecture au secondaire*. Ottawa : Conseil de recherche en sciences humaines du Canada.

Fox, E. et Riconscente, M. (2008). Metacognition and Self-Regulation in James, Piaget, and Vygotsky. *Educational Psychology Review*, *20*(4), 373-389. https://doi.org/10.1007/s10648-008-9079-2

Friedmann, N. et Reznick, J. (2021). Stages rather than ages in the acquisition of movement structures: Data from sentence repetition and 27696 spontaneous clauses. *Glossa: A Journal of General Linguistics*, *39*(1). https://doi.org/10.16995/glossa.5716

Friesen, D. C. et Bialystok, E. (2012). Metalinguistic Ability in Bilingual Children: The Role of Executive Control. *Rivista di Psicolinguistica Applicata*, *12*(3), 47-56. https://pubmed.ncbi.nlm.nih.gov/24782696, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4000604/

Fuchs, C. (2017). Pour introduire à la linguistique cognitive. Dans C. Fuchs (dir.), *La linguistique cognitive* (p. 1-24). Éditions de la Maison des sciences de l'homme.

Gabriel, A. et Poncelet, M. (2009). Validation d'une batterie de tests évaluant le niveau de conscience phonologique d'enfants d'âge préscolaire. *Glossa*, *107*, 17-38.

Gagnon, R. et Péret, C. (2016). « La phrase à deux pattes » pour favoriser l'entrée en littéracie des publics en difficulté. *Scolagram*(2). https://doi.org/https://scolagram.u-

cergy.fr/index.php?option=com_flexicontent&view=item&cid=20:deux&id=216:phrase-2-pattes&Itemid=536

Gallina, J.-M. (2011). Chapitre 01. Les représentations : un enjeu pour les sciences cognitives. Dans *Peut-on se passer de représentations en sciences cognitives ?* (p. 13-31). De Boeck Supérieur. https://doi.org/10.3917/dbu.putoi.2011.01.0011.

Garner, R. et Alexander, P. A. (1989). Metacognition: Answered and Unanswered Questions. *Educational Psychologist*, *24*(2), 143-158. https://doi.org/10.1207/s15326985ep2402_2

Gauvin, I., Marcotte, S. et Villeneuve, K. (2013). Les raisonnements grammaticaux : un outil pour développer la compétence à écrire des étudiants du postsecondaire. *Correspondance*, *19*(1).

Gauvin, I. et Thibeault, J. (2016). Pour une didactique intégrée de l'enseignement de la grammaire en contexte plurilingue québécois : le cas des constructions verbales. *Scolagram*, *2*.

Gaux, C. et Demont, E. (1997). Conscience phonologique, syntaxique et lecture. Étude chez les jeunes enfants et les pré-adolescents. Dans C. Barré-de Miniac et B. Lété (dir.), *L'illétrisme, de la prévention chez l'enfant aux stratégies de fonction chez l'adulte* (p. 203-223). De Boeck.

Gaux, C. et Gombert, J. É. (1999a). Implicit and explicit syntactic knowledge and reading in pre-adolescents. *British Journal of Developmental Psychology*, *17*, 169–188.

Gaux, C. et Gombert, J. É. (1999b). La conscience syntaxique chez les préadolescents : question de méthodes. *L'année psychologique*, *99*(1), 45-74. https://doi.org/10.3406/psy.1999.28546

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Lawrence Erlbaum Associates, Inc.

Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. Dans G. H. Bower (dir.), *The psychology of learning and motivation* (p. 217-263). Academic Press.

Gernsbacher, M. A. (1995). The Structure-Building Framework : what it is, what it might also be, and why. Dans B. K. Britton et A. C. Graesser (dir.), *Models of text understanding* (p. 289-311). Erlbaum.

Gernsbacher, M. A. (1997). Two Decades of Structure Building. *Discourse Process*, *23*(3), 265-304. https://doi.org/10.1080/01638539709544994

Gernsbacher, M. A. et Kaschak, M. P. (2013). Text Comprehension. Dans D. Reisberg (dir.), *The Oxford Handbook of Cognitive Psychology* (p. 462-474). Oxford University Press.

Gernsbacher, M. A. et Varner, K. R. (1988). *The multi-media comprehension battery (Technical Report 88-03)*. Institute of Cognitive and Decision Sciences.

Gernsbacher, M. A., Varner, K. R. et Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 430-445.

Giasson, J. (1990). *La compréhension en lecture.* Gaétan Morin Editeur.

Gineste, M.-D. et Le Ny, J.-F. (2002). *Psychologie cognitive du langage : de la reconnaissance à la compréhension*. Dunod.

Gleitman, L. R. et Rozin, P. (1977). The structure and acquisition of reading I: Relations between orthographies and the structure of language. *Toward a psychology of reading. Hillsdale, NJ: Erlbaum*, 1-53.

Goel, V., Buchel, C., Frith, C. et Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage*, *12*(5), 504-514. https://doi.org/10.1006/nimg.2000.0636

Goigoux, R. (1998). Apprendre à lire : de la pratique à la théorie. *Repères*, *18*(1), 147-162. https://doi.org/10.3406/reper.1998.2278

Goldman, S. R. et Wolfe, M. B. W. (2001). Text Comprehension: Models in Psychology. Dans N. J. Smelser et P. B. Baltes (dir.), *International Encyclopedia of the Social & Behavioral Sciences* (p. 15625-15631). Pergamon. https://doi.org/https://doi.org/10.1016/B0-08-043076-7/01542-4.

Gombert, J. É. (1988). La conscience du langage à l'âge préscolaire. *Revue française de pédagogie*(83), 65-81.

Gombert, J. É. (1990). *Le développement métalinguistique*. Presses Universitaires de France.

Gombert, J. É. (1991). Le rôle des capacités métalinguistiques dans l'acquisition de la langue écrite. *Repères*, *3*(1), 143-156. https://doi.org/10.3406/reper.1991.2023

Gombert, J. É. (1996). Activités métalinguistiques et acquisition d'une langue. *Acquisition et interaction en langue étrangère*, *8*.

Gombert, J. É. (2006). Epi/méta vs implicite/explicite: niveau de contrôle cognitif sur les traitements et apprentissage de la lecture. *Langage & pratique*, *38*, 68-76.

Gombert, J. É. (2018). Niveaux de contrôle cognitif lors des apprentissages linguistiques scolaires. *Recherches en didactique des langues et des cultures*, *15*(1). https://doi.org/10.4000/rdlc.2727

Gombert, J. É. et Fayol, M. (1988). Auto-contrôle par l'enfant de ses réalisations dans des tâches cognitives. *Revue française de pédagogie*, *82*(1), 47-59. https://doi.org/10.3406/rfp.1988.1460

Gombert, J. É., Gaux, C. et Demont, E. (1994). Capacités métalinguistiques et lecture : quels liens ? *Repères*, *9*, 61-73.

Gombert, J. É., Martinot, C. et Nocus, I. (1996). Conscience linguistique et traitement de l'information écrite. *Revue Tranel (Travaux neuchâtelois de linguistique)*, *25*, 135-153.

Goodman, K. (1989). *The Whole Language Evaluation Book*. Heinemann.

Goodwin, A. P., Petscher, Y., Carlisle, J. F. et Mitchell, A. M. (2017). Exploring the Dimensionality of Morphological Knowledge for Adolescent Readers. *Journal of Research in Reading*, *40*(1), 91-117. https://doi.org/10.1111/1467-9817.12064

Goodwin, A. P., Petscher, Y. et Reynolds, D. (2021). Unraveling Adolescent Language & Reading Comprehension: The Monster's Data. *Scientific Studies of Reading*, *26*(4), 305-326. https://doi.org/10.1080/10888438.2021.1989437

Gough, P. B. et Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, *7*(1), 6-10.

Graesser, A. C., Gernsbacher, M. A. et Goldman, R. (dir.). (2003). *Handbook of discourse processes*. Mahwah, N.J. : L. Erlbaum.

Graesser, A. C., Millis, K. K. et Zwaan, R. A. (1997). Dicourse comprehension. *Annual Review of Psychology*(48).

Graesser, A. C., Singer, M. et Trabasso, T. (1994). Constructing Inferences During Narrative Text Comprehension. *Psychological Review*, *101*(3), 371-395.

Green, L., D., McCutchen, C., Schwiebert, T., Quinlan, A., Eva-Wood, A. et Juelis, J. (2003). Morphological development in children's writing. *Journal of Educational Psychology*, *95*(4), 452–761.

Guo, Y., Roehrig, A. D. et Williams, R. S. (2011). The Relation of Morphological Awareness and Syntactic Awareness to Adults' Reading Comprehension. *Journal of Literacy Research*, *43*(2), 159-183. https://doi.org/10.1177/1086296x11403086

Hahs-Vaughn, D. L. et Lomax, R. G. (2020). *An Introduction to Statistical Concepts*. Routledge.

Hannon, B. (2012). Understanding the Relative Contributions of Lower-Level Word Processes, Higher-Level Processes, and Working Memory to Reading Comprehension Performance in Proficient Adult Readers. *Reading Research Quarterly*, *47*(2), 125–152. https://doi.org/10.1002/RRQ.013

Hannon, B. (2016). General and non-general intelligence factors simultaneously influence SAT, SAT-V, and SAT-M performance. *Intelligence*, *59*, 51-63. https://doi.org/10.1016/j.intell.2016.07.002

Hannon, B. et Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of educational psychology*, *93*(1), 103-128.

Hannon, B. et Daneman, M. (2006). What do tests of reading comprehension ability such as the VSAT really measure?: a componential analysis. Dans A. V. Mitel (dir.), *Focus on Educational Psychology* (p. 105-146). Nova Science Publishers, Inc.

Hauser, M. D., Chomsky, N. et Fitch, T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, *298*, 1569-1579.

Heine, J. H., Gebhard, M., Schwab, S., Neumann, P., Gorges, J. et Elke, W. (2018). Testing psychometric properties of the CFT 1- R for students with special educational needs. *Psychological Test and Assessment Modeling*, *60*(1), 3-27.

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. http://dx.doi.org/10.1080/10824660802427710

Hermanto, N., Moreno, S. et Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: How bilingual? *International*

Journal of Bilingual Education and Bilingualism, *15*(2), 131-145. https://doi.org/10.1080/13670050.2011.652591

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 G. B. Thompson, W. E. Tunmer et T. Nicholson (dir.), *Reading acquisition processes*.

Houdé, O. (2023). Chapitre II. Les stades de l'intelligence chez l'enfant selon Jean Piaget. Dans O. Houdé (dir.), *Comment raisonne notre cerveau* (p. 358-384). Presses Universitaires de France.

Huot, D. et Schmidt, R. (1996). Conscience et activité métalinguistique. Quelques points de rencontre. *Acquisition et interaction en langue étrangère [En ligne]*(8). http://journals.openedition.org/aile/1237

Hwaszcz, K. et Klimek-Jankowska, D. (2017). The Role of Semantic Transparency in the Processing of Compound Words in Polish: Evidence from a Masked Priming Experiment. *Studies in Polish Linguistics*, *12*(3), 145-171.

Hyams, N. et Orfitelli, R. (2015). The Acquisition of Syntax. Dans H. Cairns et E. Fernandez (dir.), *Handbook of Psycholinguistics*. Wiley/Blackwell Publishers.

Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S. et Zorman, M. (2010). *BALE : Batterie Analytique du Langage Ecrit*. UPMF.

Jeon, E. H. (2011). Contribution of Morphological Awareness to Second-Language Reading Comprehension. *The Modern Language Journal*, *95*(2), 217-235. https://doi.org/10.1111/j.1540-4781.2011.01179.x

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, *107*(43), 18243-18250. https://doi.org/10.1073/pnas.1012933107

Johnson-Laird, P. N. (2013). The Mental Models Perspective. Dans D. Reisberg (dir.), *The Oxford Handbook of Cognitive Psychology*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195376746.013.0041.

Johnson-Laird, P. N. et Khemlani, S. S. (2014). Toward a Unified Theory of Reasoning. *Psychology of Learning and Motivation*, *59*, 1-42. https://doi.org/10.1016/b978-0-12-407187-2.00001-0

Johnson-Laird, P. N., Khemlani, S. S. et Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, *19*(4), 201-214. https://doi.org/10.1016/j.tics.2015.02.006

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.

Kahneman, D. et Frederick, S. (2002). Representativeness Revisited: Attribute Substitution in Intuitive Judgment. Dans T. Gilovich, D. Griffin et D. Kahneman (dir.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (p. 49-81). Cambridge University Press. https://doi.org/doi:10.1017/CBO9780511808098.004.

Kail, M. (2000). Chapitre 1. Acquisition syntaxique et diversité linguistique. Dans *L'acquisition du langage. Vol. II* (p. 9-44). https://doi.org/10.3917/puf.fayol.2000.02.0009.

Kaput, J. J., Carraher, D. W. et Blanton, M. L. (2008). *Algebra in the Early Grades*. Taylor & Francis.

Karimi, H. et Ferreira, F. (2016). Good-enough linguistic representations and online cognitive equilibrium in language processing. *Quarterly Journal of Experimental Psychology*, *69*(5), 1013-1040. https://doi.org/10.1080/17470218.2015.1053951

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, *23*(2).

Keenan, J. M., Betjemann, R. S. et Olson, R. K. (2008). Reading Comprehension Tests Vary in the Skills They Assess: Differential Dependence on Decoding and Oral Comprehension. *Scientific Studies of Reading*, *12*(3), 281-300. https://doi.org/10.1080/10888430802132279

Kendeou, P., McMaster, K. L. et Christ, T. J. (2016). Reading Comprehension. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 62-69. https://doi.org/10.1177/2372732215624707

Khomsi, A. (1990). *Epreuve d'évaluation de la compétence en lecture : lecture de mots et compréhension*. Les éditions du Centre de Psychologie Appliquée.

Khomsi, A. (1999). *Lecture de mots et compréhension, revisée*. Éditions du Centre de psychologie appliquée.

Kieffer, M. J. et Lesaux, N. K. (2012). Direct and Indirect Roles of Morphological Awareness in the English Reading Comprehension of Native English, Spanish, Filipino, and Vietnamese Speakers. *Language Learning*, *62*(4), 1170-1204. https://doi.org/10.1111/j.1467-9922.2012.00722.x

Kieffer, M. J., Petscher, Y., Proctor, C. P. et Silverman, R. D. (2016). Is the Whole Greater Than the Sum of Its Parts? Modeling the Contributions of Language Comprehension Skills to Reading Comprehension in the Upper Elementary Grades. *Scientific Studies of Reading*, 1-19.

Kim, Y. S. G., Guo, Q., Liu, Y., Peng, Y. et Yang, L. (2019). Multiple Pathways by Which Compounding Morphological Awareness Is Related to Reading Comprehension: Evidence From Chinese Second Graders. *Reading Research Quarterly*, *55*(2), 193-212. https://doi.org/10.1002/rrq.262

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kintsch, W. et van Dijk, T. A. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review*, *85*(5), 363-394.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L. et Parrila, R. (2011). Children's morphological awareness and reading ability. *Reading and Writing: An Interdisciplinary Journal*, *25*(2), 389-410. https://doi.org/10.1007/s11145-010-9276-5

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3 éd.). The Guilford Press.

Kunnan, A. J. (1998). An introduction to structural equation modelling for language assessment research. *Language Testing*, *13*(3), 295–332.

Kuo, L.-j. et Anderson, R. C. (2006). Morphological Awareness and Learning to Read: A Cross-Language Perspective. *Educational Psychologist*, *41*(3), 161-180. https://doi.org/10.1207/s15326985ep4103_3

Lavric, A., Rastle, K. et Clapp, A. (2011). What do fully visible primes and brain potentials reveal about morphological decomposition? *Psychophysiology*, *48*(5), 676-686. https://doi.org/10.1111/j.1469-8986.2010.01125.x

Law, J. M., Wouters, J. et Ghesquiere, P. (2015). Morphological Awareness and Its Role in Compensation in Adults with Dyslexia. *Dyslexia*, *21*(3), 254-272. https://doi.org/10.1002/dys.1495

Le Ny, J.-F. (2017). Éléments de psycholinguistique cognitive : des représentations à la compréhension. Dans C. Fuchs (dir.), *La linguistique cognitive* (p. 155-170). Éditions de la Maison des sciences de l'homme.

Lecocq, P. (1996). *L'ECOSSE: une épreuve de compréhension syntaxico-sémantique*. Presses Universitaires du Septentrion.

Leontiev, A. N. (1984). *Activité, conscience, personnalité*. Moscou Editions du Progrès.

Lepoire Duc, S. et Sautot, J.-P. (2010). *Expliquer la grammaire : Comprendre les concepts de la langue*. Scéren/ CRDP Académie de Grenoble. https://halshs.archives-ouvertes.fr/halshs-00986685

Lesaux, N. K. et Siegel, L. S. (2003). The Development of Reading in Children Who Speak English as a Second Language. *Developmental Psychology*, *39*(6), 1005-1019.

Levesque, K. C., Breadmore, H. L. et Deacon, S. H. (2020). How morphology impacts reading and spelling: advancing the role of morphology in models of literacy development. *Journal of Research in Reading*, *44*(1), 10-26. https://doi.org/10.1111/1467-9817.12313

Levesque, K. C., Kieffer, M. J. et Deacon, S. H. (2017). Morphological awareness and reading comprehension: Examining mediating factors. *Journal of Experimental Child Psychology*, *160*, 1-20. https://doi.org/10.1016/j.jecp.2017.02.015

Li, L. et Wu, X. (2015). Effects of metalinguistic awareness on reading comprehension and the mediator role of reading fluency from grades 2 to 4. *PLoS One*, *10*(3), e0114417. https://doi.org/10.1371/journal.pone.0114417

Loehlin, J. C. (2003). *Latent variable model. An introduction to factor, path, and structural equation analysis* (4 éd.). Psychology Press.

Lonigan, C. J. (2015). Literacy Development. Dans R. M. Lerner (dir.), *Handbook of Child Psychology and Developmental Science* (p. 1-43). John Wiley & Sons.

MacGinitie, W. H. et MacGinitie, R. K. (1989). *Gates-MacGinitie Reading Tests*. Riverside.

MacGinitie, W. H., MacGinitie, R. K., Maria, K. et Dreyer, L. G. (2000). *Gates-MacGinitie Reading test* (4 éd.). Riverside Publishing Company.

MacKay, E., Lynch, E., Sorenson Duncan, T. et Deacon, S. H. (2021). Informing the Science of Reading: Students' Awareness of Sentence ‐ Level Information Is Important for Reading Comprehension. *Reading Research Quarterly*, *56*(S1). https://doi.org/10.1002/rrq.397

Mackintosh, N. J. et Bennett, E. S. (2005). What do Raven's Matrices measure? An analysis in terms of sex differences. *Intelligence*, *33*(6), 663-674. https://doi.org/10.1016/j.intell.2005.03.004

Mahony, D., Singson, M. et Mann, V. (2000). *Reading and Writing: An Interdisciplinary Journal*, *12*(3/4), 191-218. https://doi.org/10.1023/a:1008136012492

Malakoff, M. et Hakuta, K. (1991). Translation skill and metalinguistic awareness in bilinguals. Dans E. Bialystok (dir.), *Language Processing in Bilingual Children* (p. 141-166). Cambridge University Press.

Manolitsis, G., Grigorakis, I. et Georgiou, G. K. (2017). The Longitudinal Contribution of Early Morphological Awareness Skills to Reading Fluency and Comprehension in Greek. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01793

Marcotte, F. (2021). *Contribution de la conscience syntaxique à la compréhension en lecture en langue maternelle et en langue seconde : étude des effets médiateurs des habiletés syntaxiques et de la reconnaissances de mots* [Mémoire de maîtrise, Université du Québec à Montréal]. Montréal.

Marinier, J. P. et Laird, J. E. (2006). *A Cognitive Architecture Theory of Comprehension and Appraisal*. n.a.

Marmy Cusin, V., Gagnon, R., Darme-Xu, A. et Bulea Bronckart, E. (2022). Pratiques courantes de l'enseignement grammatical : les cas des objets complément de nom et valeurs des temps du passé. *Congrès Mondial de Linguistique Française. SHS Web of Conferences*, *138*. https://doi.org/10.1051/shsconf/202213806017

McNamara, D. et Magliano, J. P. (2009). Toward a Comprehensive Model of Comprehension. *Psychology of Learning and Motivation*, *51*, 297-384. https://doi.org/10.1016/S0079-7421(09)51009-2

Mehmetoglu, M. et Venturini, S. (2021). *Structural Equation Modelling with Partial Least Squares Using Stata and R*. CRC Press.

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. http://dx.doi.org/10.1080/00313831.2011.594611

MELS. (2009). *Progression des apprentissages au primaire : français langue d'enseignement*. Québec : Gouvernement du Québec.

Mercier, H. et Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*(2), 57-74; discussion 74-111. https://doi.org/10.1017/S0140525X10000968

Millis, K. et Larson, M. (2008). Applying the Construction-Integration Framework to Aesthetic Responses to Representational Artworks. *Discourse Processes*, *45*(3), 263-287. https://doi.org/10.1080/01638530802070015

Morgenstern, A., Parisse, C. et (de) Pontonx, S. (2014). « On dit pas je veux ! » Apprentissage explicite et implicite du conditionnel dans les interactions adulte-enfant. *Langage, Interaction et Acquisition /Language, Interaction and Acquisition*(1), 19-37.

Muse, A. E. (2005). *The nature of morpholological knowledge* [Thèse de doctorat, Florida State University]. Tallahassee.

Nadeau, M. (2017). L'implantation d'une nouvelle grammaire pédagogique au Québec : obstacles et opportunités. *Caplletra*(63), 89-216.

Nadeau, M. et Fisher, C. (2006). *La grammaire nouvelle - La comprendre et l'enseigner*.

Nadeau, M. et Fisher, C. (2011). Les connaissances implicites et explicites en grammaire : quelle importance pour l'enseignement? Quelles conséquences? *Bellaterra Journal of Teaching & Learning Language & Literature*, *4*(4). https://doi.org/10.5565/rev/jtl3.446

Nagy, W. E. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. Dans *Vocabulary acquisition: Implications for reading comprehension.* (p. 52-77). Guilford Press.

Nagy, W. E., Berninger, V., Abbott, R., Vaughan, K. et Vermeulen, K. (2003). Relationship of Morphology and Other Language Skills to Literacy Skills in At-Risk Second-Grade Readers and At-Risk Fourth-Grade Writers. *Journal of Educational Psychology*, *95*(4), 730-742. https://doi.org/10.1037/0022-0663.95.4.730

Nagy, W. E., Berninger, V. W. et Abbott, R. D. (2006). Contributions of Morphology Beyond Phonology to Literacy Outcomes of Upper Elementary and Middle-School Students. *Journal of Educational Psychology*, *98*(1), 134-147. https://doi.org/10.1037/0022-0663.98.1.134

Nation, K. et Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied Psycholinguistics*, *21*(2), 229-241. https://doi.org/10.1017/S0142716400002046

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S. S., Naccache, L., Hale, J. T., Pallier, C. et Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, *114*(18), E3669-E3678. https://doi.org/10.1073/pnas.1701590114

O'Reilly, R. C. et Munakata, Y. (2000). *Computational explorations in cognitive neuroscience*. MIT Press.

Oakhill, J. et Garnham, A. (1988). *Becoming a skilled reader*. Basil Blackwell.

OCDE. (2001). *Connaissances et compétences : des atouts pour la vie. Premiers résultats du programme international de l'OCDE pour le suivi des acquis des élèves (PISA) 2000*. Les éditions de l'OCDE.

OCDE. (2014). *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.*

Ockey, G. et Choi, I. (2015). Structural Equation Modeling Reporting Practices for Language Assessment. *Language Assessment Quarterly*, 305-319.

Pan, J., Song, S., Su, M., McBride, C., Liu, H., Zhang, Y., Li, H. et Shu, H. (2015). On the relationship between phonological awareness, morphological awareness and Chinese literacy skills: evidence from an 8-year longitudinal study. *Developmental Science*. https://doi.org/10.1111/desc.12356

Park, H. R. et Uno, A. (2015). Cognitive Abilities Underlying Reading Accuracy, Fluency and Spelling Acquisition in Korean Hangul Learners from Grades 1 to 4: A Cross-Sectional Study. *Dyslexia*, *21*(3), 235-253. https://doi.org/10.1002/dys.1500

Parsons, L. M. et Osherson, D. (2001). New Evidence for Distinct Right and Left Brain Systems for Deductive versus Probabilistic Reasoning. *Cerebral cortex*, *11*(10), 954-965.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Hochenberger, R., Sogo, H., Kastman, E. et Lindelov, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195-203. https://doi.org/10.3758/s13428-018-01193-y

Perfetti, C. A. (1985). *Reading ability*. Oxford University Press.

Perfetti, C. A., Landi, N. et Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. Dans M. J. Snowling et C. Hulme (dir.), *The Science of Reading: A Handbook*. https://doi.org/https://doi.org/10.1002/9780470757642.ch13.

Perfetti, C. A. et Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences.

Piaget, J. (1947). *La psychologie de l'intelligence*. Colin.

Piaget, J. (1974). *Réussir et comprendre*. Presses Universitaires de France.

Pinker, S. et Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition*, *95*(2), 201-236. https://doi.org/10.1016/j.cognition.2004.08.004

Pintrich, P. R. (2002). The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing. *Theory Into Practice*, *41*(4), 219-225.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S. et Patterson, K. (1996). Understanding Normal and Impaired Word Reading: Computational Principles in Quasi-Regular Domains. *Psychological Review*, *103*(1), 56-115.

Plaza, M. et Cohen, H. (2003). The interaction between phonological processing, syntactic awareness, and naming speed in the reading and spelling performance of first-grade children. *Brain and Cognition*, *53*(2), 287-292. https://doi.org/10.1016/s0278-2626(03)00128-3

Pratt, C., Tunmer, W. E. et Bowey, J. A. (1984). Children's capacity to correct grammatical violations in sentences. *Journal of Child Language*, *11*(1), 129-141. https://doi.org/10.1017/S0305000900005626

Quevillon Lacasse, C., Nadeau, M. et Giguère, M.-H. (2018). La combinaison de phrases : un dispositif stimulant et efficace pour développer la créativité syntaxique. *Phrases*, *23*(5).

R Core Team. (2020). *R: A language and environment for statistical computing*. (version https://www.R-project.org/

Rastle, K. (2009). Visual word recognition. Dans M. Gaskell (dir.), *The Oxford Handbook of psycholinguistics*. Oxford University Press.

Raven, J. (1938). *Guide to Using Progressive Matrices (1938): The Design and Use of the Scale-Instructions-Records and Marking-Key-Normal Score Composition-Norms for Children and for Adults Up to the Age of 65-Bibliography*. HK Lewis.

Ravid, D. et Tolchinsky, L. (2002). Developing linguistic literacy: a comprehensive model. *Journal of Child Language*, *29*(2), 417-447. https://doi.org/10.1017/s0305000902005111

Reder, F., Marec-Breton, N., Gombert, J. É. et Demont, E. (2013). Second-language learners' advantage in metalinguistic awareness: a question of languages' characteristics. *British Journal of Educational Psychology*, *83*(4), 686-702. https://doi.org/10.1111/bjep.12003

Redmond, L., Bourdages, R. et Foucambert, D. (2021). La conscience morphologique en français langue seconde : dimensionnalité et implications didactiques dans les programmes d'immersion précoce. *Revue de l'AQEFLS*, *34*(2). https://doi.org/https://doi.org/10.7202/1082347ar

Reuter, Y., Cohen-Azria, C., Daunay, B., Delcambre, I. et Lahanier-Reuter, D. (2013). Métalangage — activité métalinguistique. Dans *Dictionnaire des concepts fondamentaux des didactiques* (p. 123-128). De Boeck Supérieur. https://www.cairn.info/dictionnaire-des-concepts-fondamentaux-des-didacti--9782804169107-page-123.htm.

Riegel, M., Pellat, J.-C. et Rioul, R. (2004). *Grammaire méthodique du français*. Presses Universitaires de France.

Roberts, R. et Gibson, E. (2002). Individual Differences in Sentence Memory. *Journal of Psycholinguistic Research*, *31*(6), 573-598.

Rock, I. et Palmer, S. (1990). The legacy of Gestalt psychology. *Scientific American*, *263*(6), 84-90. https://doi.org/10.1038/scientificamerican1290-84

Roegiers, X. (2007). Situation, problème, situation-problème. Dans *Des situations pour intégrer les acquis scolaires* (p. 15-76).

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Roy, C. et Labelle, M. (2007). Connaissance de la morphologie dérivationnelle chez les francophones et non-francophones de 6 à 8 ans. *Revue canadienne de linguistique appliquée*, 263-291.

Rubin, H. (1988). Morphologicial Knowledge and Early Writing Ability. *Language and Speech*, *31*(4), 337-355.

Rupp, A. A. (2012). Psychological versus psychometric dimensionality in reading assessment. Dans J. P. Sabatini, E. Albro et T. O'Reilly (dir.), *Measuring up: Advances in how to assess reading ability* (p. 135-151). Rowman & Littlefield Education.

Rupp, A. A., Ferne, T. et Choi, H. (2016). How assessing reading comprehension with multiple-choice questions shapes the construct: a cognitive processing perspective. *Language Testing*, *23*(4), 441-474. https://doi.org/10.1191/0265532206lt337oa

Saffran, J. R. (2003). Statistical Language Learning: Mechanisms and Constraints. *American Psychological Society*, *12*(4), 110-114.

Sander, E. (2016). Enjeux sémantiques pour les apprentissages arithmétiques. *Bulletin de psychologie*, *Numéro 546*(6), 463-469. https://doi.org/10.3917/bupsy.546.0463

Saunders, J. A., Morrow-Howell, N., Spitznagel, E., Doré, P., Proctor, E. K. et Pescarino, R. (2006). Imputing Missing Data: A Comparison of Methods for Social Work Researchers. *Social Work Research*, *31*(1), 20-31.

Sawyer, D. J. et Fox, B. J. (1991). *Phonological awareness in reading: The evolution of current perspectives*. Springer-Verlag Publishing. https://doi.org/10.1007/978-1-4612-3010-6

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M. et Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology*, *91*(3), 439-449. https://doi.org/10.1037/0022-0663.91.3.439

Schönefeld, D. (2011). *Where Lexicon and Syntax meet*. De Gruyter Mouton. https://doi.org/doi:10.1515/9783110849066.89

Seidenberg, M. S. et McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review*, *96*(4), 523-568.

Shea, N. et Frith, C. D. (2019). The Global Workspace Needs Metacognition. *Trends in Cognitive Sciences*, *23*(7), 560-571. https://doi.org/10.1016/j.tics.2019.04.007

Shi, R. (2014). Functional Morphemes and Early Language Acquisition. *Child Development Perspectives*, *8*(1), 6-11. https://doi.org/10.1111/cdep.12052

Simard, C. et Chartrand, S.-G. (2011). *Grammaire de base : 2e et 3e cycle du primaire*. Éditions du Renouveau pédagogique.

Simard, C., Dufays, J.-L., Dolz, J. et Garcia-Debanc, C. (2010). Qu'est-ce que la didactique du français ? Quelques notions préliminaires. Dans *Didactique du français langue première* (p. 9-39). De Boeck Supérieur.

Simard, D. (2008). Effet de la nature des éléments grammaticaux sur la saisie en français langue seconde. *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, *64*, 491-522.

Simard, D., Fortier, V. et Foucambert, D. (2011). Measuring metasyntactic ability among heritage language children. *Bilingualism: Language and Cognition*, *16*(01), 19-31.

Simard, D., Foucambert, D. et Labelle, M. (2012a). Examining the contribution of metasyntactic ability to reading comprehension among native and non-native speakers of French. *International Journal of Bilingualism*.

Simard, D., Foucambert, D. et Labelle, M. (2012b). Examining the contribution of metasyntactic ability to reading comprehension among native and non-native speakers of French. *International Journal of Bilingualism*, *18*(6), 586-604. https://doi.org/10.1177/1367006912452169

Simard, D., French, L. et Fortier, V. (2007). Elicited metalinguistic reflection and second language learning: Is there a link? *System*, *35*(4), 509-522. https://doi.org/10.1016/j.system.2007.06.004

Simard, D., Labelle, M. et Bergeron, A. (2017). Measuring Metasyntactic Abilities: On a Classification of Metasyntactic Tasks. *Journal of Psycholinguistic Research*, *46*(2), 433-456. https://doi.org/10.1007/s10936-016-9445-z

Simpson, I. C., Moreno-Pérez, F. J., Rodríguez-Ortiz, I. d. l. R., Valdés-Coronel, M. et Saldaña, D. (2019). The effects of morphological and syntactic knowledge on reading comprehension in spanish speaking children. *Reading and Writing*, *33*(2), 329-348. https://doi.org/10.1007/s11145-019-09964-5

Singson, M., Mahony, D. et Mann, V. (2000). *Reading and Writing*, *12*(3/4), 219-252. https://doi.org/10.1023/a:1008196330239

Snow, C. E., Burns, S. M. et Griffin, P. (1998). *Preventing reading difficulties in young children*. National Academy Press.

Son, L. K. et Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. Dans B. L. Schwartz et T. J. Perfect (dir.), *Applied Metacognition* (p. 15-38). Cambridge University Press. https://doi.org/10.1017/CBO9780511489976.003.

Song, Y., Do, Y., Lee, J., Thompson, A. L. et Waegemaekers, E. R. (2019). The reality of hierarchical morphological structure in multimorphemic words. *Cognition*, *183*, 269-276. https://doi.org/10.1016/j.cognition.2018.10.015

Spearman, C. E. (1904). 'General intelligence,' objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201-293. https://doi.org/10.2307/1412107

Spearman, C. E. (1946). Theory of the general factor. *British Journal of Psychology*, *36*, 117–131.

Spencer, M., Muse, A., Wagner, R. K., Foorman, B., Petscher, Y., Schatschneider, C., Tighe, E. L. et Bishop, M. D. (2015). Examining the underlying dimensions of morphological awareness and vocabulary knowledge. *Reading and Writing*, *28*(7), 959-988. https://doi.org/10.1007/s11145-015-9557-0

Spencer, M. et Wagner, R. K. (2018). The Comprehension Problems of Children with Poor Reading Comprehension despite Adequate Decoding: A Meta-Analysis. *Review of Educational Research*, *88*(3), 366-400. https://doi.org/10.3102/0034654317749187

Stanovich, K. E., Cunningham, A. E. et Cramer, B. B. (1984). Assessing Phonological Awareness in Kindergarten Children: Issues of Task Comparability. *Jounal of Experimental Child Psychology*, *38*, 175-190.

Swanson, H. L., Harris, K. R. et Graham, S. (2013). *Handbook of learning disabilities*. http://site.ebrary.com/id/10648581

Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behavior*, *20*(3), 289-297. https://doi.org/https://doi.org/10.1016/S0022-5371(81)90439-4

The jamovi project. (2021). *jamovi (Version 1.6)*. (version https://www.jamovi.org

Thievenaz, J. (2015). L'étonnement, ressource pour l'apprentissage au travail et en formation. *Travail et Apprentissages*, *N° 17*(1), 81-102. https://doi.org/10.3917/ta.017.0081

Thompson, V. A., Prowse Turner, J. A. et Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107-140. https://doi.org/10.1016/j.cogpsych.2011.06.001

Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, *8*, 323–332. https://doi.org/https://doi.org/10.1037/h0075325

Tibi, S. et Kirby, J. R. (2017). Morphological awareness: Construct and predictive validity in Arabic. *Applied Psycholinguistics*, *38*(05), 1019-1043. https://doi.org/10.1017/s0142716417000029

Tighe, E. L. (2015). *Assessing the Importance of Metalinguistic Skills to the Word Reading and Reading Comprehension Abilities of Adult Basic Education Students.* [Thèse de doctorat, Florida State University].

Tighe, E. L. et Binder, K. S. (2015). An investigation of morphological awareness and processing in adults with low literacy. *Applied Psycholinguistics*, *36*(2), 245-273. https://doi.org/10.1017/S0142716413000222

Tighe, E. L., Little, C. W., Arrastia-Chisholm, M. C., Schatschneider, C., Diehm, E., Quinn, J. M. et Edwards, A. A. (2018). Assessing the direct and indirect effects of metalinguistic awareness to the reading comprehension skills of struggling adult readers. *Reading and Writing*. https://doi.org/10.1007/s11145-018-9881-2

Tighe, E. L. et Schatschneider, C. (2015). Exploring the Dimensionality of Morphological Awareness and Its Relations to Vocabulary Knowledge in Adult Basic Education Students. *Reading Research Quarterly*, *50*(3), 293-311. https://doi.org/10.1002/rrq.102

Tighe, E. L. et Schatschneider, C. (2016). Modeling the Relations Among Morphological Awareness Dimensions, Vocabulary Knowledge, and Reading Comprehension in Adult Basic Education Students. *Frontiers in Psychology*, *7*, 86. https://doi.org/10.3389/fpsyg.2016.00086

To, N. L., Tighe, E. L. et Binder, K. S. (2016). Investigating morphological awareness and the processing of transparent and opaque words in adults with low literacy skills and in skilled readers. *Journal of Research in Reading*, *39*(2), 171-188. https://doi.org/10.1111/1467-9817.12036

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, *11*(1), 61-82.

Tong, X., Deacon, S. H. et Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: another piece of the puzzle. *Journal of Learning Disabilities*, *47*(1), 22-33. https://doi.org/10.1177/0022219413509971

Trabasso, T., van den Broek, P. et Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, *12*(1), 1-25. https://doi.org/10.1080/01638538909544717

Treimen, R. et Kessler, B. (2009). Learning to read. Dans M. Gaskell (dir.), *The Oxford handbook of Psycholinguistics*. Oxford University Press.

Tunmer, W. E. (1989a). Conscience phonologique et acquisition de la langue écrite. Dans L. Rieben et C. Perfetti (dir.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques*.

Tunmer, W. E. (1989b). The role of language-related factors in reading disability. Dans *Phonology and reading disability: Solving the reading puzzle.* (p. 91-131). The University of Michigan Press.

Tunmer, W. E. (1990). The role of language prediction skills in beginning reading. *New Zealand Journal of Educational Studies*, *25*(2), 95-114.

Tunmer, W. E. (1997). *Metalinguistic Skills in Reading Development Encyclopedia of Language and Education* (pp. 27–36).

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. http://www.ncbi.nlm.nih.gov/pubmed/22293683

Tunmer, W. E. et Herriman, M. L. (1984). The Development of Metalinguistic Awareness: A Conceptual Overview. Dans W. E. Tunmer, C. Pratt et M. L. Herriman (dir.), *Metalinguistic Awareness in Children: Theory, Research, and Implications* (p. 12-35). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-69113-3_2.

Tyler, A. et Nagy, W. E. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language*, *28*(6), 649–667.

van Dijk, T. A. et Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press.

Vaughn, S., Roberts, G., Wexler, J., Vaughn, M. G., Fall, A. M. et Schnakenberg, J. B. (2015). High School Students With Reading Comprehension Difficulties: Results of a Randomized Control Trial of a Two-Year Reading Intervention. *Journal of Learning Disabilities*, *48*(5), 546-558. https://doi.org/10.1177/0022219413515511

Verhoeven, L. et Perfetti, C. (2008). Advances in text comprehension: model, process and development. *Applied Cognitive Psychology*, *22*(3), 293-301. https://doi.org/10.1002/acp.1417

Vygotsky, L. S. (1930/1933/1978). *Mind in society: The development of higher mental processes* : Cambridge, MA: Harvard University Press. (Oeuvre originale publiée en 1930, 1933).

Vygotsky, L. S. (1934/1997). *Pensée et langage*. La Dispute.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A. et Pearson, N. A. (2010). *TOSREC: Test of silent reading efficiency and comprehension*. Pro-Ed.

Wechsler, D. (2016). *Echelle d'Intelligence de Wechsler pour Enfants et Adolescents – 5ème édition*. ECPA.

Weiß, R. H. et Osterland, J. (2013). *CFT 1-R Grundintelligenztest Skala 1 - Revision*. Hogrefe.

Wilson-Fowler, E. B. et Apel, K. (2015). Influence of Morphological Awareness on College Students' Literacy Skills. *Journal of Literacy Research*, *47*(3), 405-432. https://doi.org/10.1177/1086296x15619730

Woll, N. (2019). How French speakers reflect on their language: a critical look at the concept of metalinguistic awareness. *Language Awareness*, *28*(1), 49-73. https://doi.org/10.1080/09658416.2019.1567743

Woodcock, R. W. (1987). *Woodcock reading mastery tests—Revised*. American Guidance Service.

Woodcock, R. W., McGrew, K. S. et Mather, N. (2001). *Woodcock-Johnson III*. Riverside.

Xuelian, Z., Raquel, M. et Aryadoust, V. (2019). Structural equation modeling to predict performance in English proficiency tests. Dans V. Aryadoust et M. Raquel (dir.), *Quantitative Data Analysis for Language Assessment Volume II - Advanced Methods* (vol. II). Routledge.

Yanuar, F., Devianto, D., Marisa, S. et Zetra, A. (2015). Consistency test of reliability index in SEM model. *Applied Mathematical Sciences*, *9*, 5283-5292. https://doi.org/10.12988/ams.2015.56446

Zhang, J., Anderson, R. C., Wang, Q., Packard, J., Wu, X., Tang, S. et Ke, X. (2011). Insight into the structure of compound words among speakers of Chinese and English. *Applied Psycholinguistics*, *33*(4), 753-779. https://doi.org/10.1017/s0142716411000555

Zhang, L. (2018). *Metacognitive and Cognitive Strategy Use in Reading Comprehension. A Structural Equation Modelling Approach*. Springer.

Zinbarg, R. E., Revelle, W., Yovel, I. et Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123-133. https://doi.org/10.1007/s11336-003-0974-7

Zorman, M., Bressoux, P., Bianco, M., Lequette, C., Pouget, G. et Pourchet, M. (2015). « PARLER » : un dispositif pour prévenir les difficultés scolaires. *Revue française de pédagogie*(193), 57-76. https://doi.org/10.4000/rfp.4890

Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes*, *28*(1), 81-88. https://doi.org/10.1080/01638539909545070

Zwaan, R. A. (2001). Situation Model: Psychological. Dans N. J. Smelser et P. B. Baltes (dir.), *International Encyclopedia of the Social & Behavioral Sciences* (p. 14137-14141). Pergamon. https://doi.org/https://doi.org/10.1016/B0-08-043076-7/01550-3.

Zwaan, R. A., Langston, M. C. et Graesser, A. C. (1995). The construction of situation models in narrative comprehension. *Psychological Science*, *6*(5), 292-297.

Zwaan, R. A., Magliano, J. P. et Graesser, A. C. (1995). Dimensions of Situation Model Construction in Narrative Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 386-397.