De l'automaticité à l'expressivité et à la compréhension en lecture : Évaluation et développement de la prosodie en lecture chez le jeune lecteur

Erika Godde

▶ **To cite this version:**

Erika Godde. De l'automaticité à l'expressivité et à la compréhension en lecture : Évaluation et développement de la prosodie en lecture chez le jeune lecteur. Psychologie. Université Grenoble Alpes [2020-..], 2020. Français. NNT : 2020GRALS009 . tel-03150403v2

HAL Id: tel-03150403

https://theses.hal.science/tel-03150403v2

Submitted on 9 Mar 2021 (v2), last revised 24 Feb 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : PCN - Sciences cognitives, psychologie et neurocognition

Arrêté ministériel : 25 mai 2016

Présentée par

Erika GODDE RASSAT

Thèse dirigée par **Marie-Line BOSSE, Pr, LPNC, Université Grenoble Alpes et codirigée par Gérard BAILLY , DR, GIPSA-Lab, CNRS**

préparée au sein des **Laboratoires LPNC et GIPSA-Lab**
dans **l'École Doctorale Ingénierie pour la Santé la Cognition et l'Environnement**

De l'automaticité à l'expressivité et à la compréhension en lecture : Évaluation et développement de la prosodie en lecture chez le jeune lecteur

From automaticity to expressivity and comprehnsion in reading :Assessment and development of reading prosody in young readers

Thèse soutenue publiquement le **27 novembre 2020**,
devant le jury composé de :

Madame MARIE-LINE BOSSE
PROFESSEUR DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES, Directrice de thèse
Monsieur ERIC LAMBERT
PROFESSEUR DES UNIVERSITES, UNIVERSITE DE POITIERS, Rapporteur
Madame ELISABETH DELAIS-ROUSSARIE
DIRECTRICE DE RECHERCHE, LLING, CNRS, Rapporteure
Monsieur GERARD BAILLY
DIRECTEUR DE RECHERCHE, CNRS DELEGATION ALPES, Co-directeur de thèse
Monsieur EDOUARD GENTAZ
PROFESSEUR, UNIVERSITE DE GENEVE - SUISSE, Président
Madame ANNA POTOCKI
MAITRE DE CONFERENCES, UNIVERSITE GRENOBLE ALPES, Examinatrice

Remerciements

Je remercie évidemment en premier lieu mes directeurs de thèse pour m'avoir accompagnée, encouragée et guidée pendant ces trois ans et demi. Vous avez réussi à me laisser toute l'autonomie dont j'avais besoin tout en étant toujours présents pour répondre à mes questions. Je sais à quel point cet équilibre peut-être difficile et j'ai vraiment apprécié de travailler avec vous. Merci à Marie-Line d'avoir patiemment surmonté les fautes d'orthographe dans la relecture de ce manuscrit. Merci à Gérard d'avoir patiemment lu mes grandes phrases et tout ce que j'avais à dire.

Merci ensuite au jury d'avoir d'avoir accepté d'examiner ce travail. Edouard Gentaz et Anna Potocki d'avoir accepté d'être examinateurs. Elizabeth Delais-Roussarie et Eric Lambert d'avoir accepté le rôle de rapporteurs.

Merci aux directrices des écoles, Nathalie Michel, Marianne Salvete et Nathalie Carpentier, et au principal du collège, Mr Piroud. Merci à tous les enseignants qui ont acccepté de se laisser embarquer dans le projet, de me recevoir dans leurs classes, d'ajuster leurs emplois du temps : Mmes Robin et Monteagudot au collège, Agnès Fénérol, Sophie Lopez, Yann Janvier, Véronique Rochas, Mme Caponi, Christophe Fernandez, Florence Durand, Nathalie Caroff, Caroline Filiputti, Betty Charensol, Juliette Dos Santos, Marianne Dagnaux et tous leurs remplaçants. Merci de votre acceuil, votre intérêt pour ce travail et pour tous les échanges trés intéressants que nous avons eu à ce sujet.

Merci surtout à tous les enfants qui ont participé aux enregistrements et que j'ai retrouvés régulièrement avec plaisir au long de ces trois ans. Ce sont certainement les personnes les plus importantes dans ce travail de thèse. Ils ont tous été à l'écoute, intéressés et plein de questions sur le matériel et ce qu'ils voyaient à l'écran. Ils ont tous essayé de ne pas trop gigoter avec les ceintures, de ne pas trop crier dans le micro, de ne pas courir dans les escaliers et même parfois d'empiéter un peu sur la récré pour rester avec moi. Et surtout ils ont tous mis du coeur dans leur lecture. Un grand merci à eux !

Merci aux stagiaires qui sont venues m'assister dans la récolte des données, qui ont appris à serrer les ceintures respiratoires suffisamment, mais pas trop, et à répondre aux questions des enfants, sans trop leur en dire. Elles ont enregistré, corrigé des alignements et évalué des centaines de lectures toujours avec le sourire ! Merci à Manon Metz, Anne-Laure Dugué et Camille Deshayes, futures orthophonistes, à Laurence Vallin et Liliane Limbert qui, comme moi se posaient des questions les pratiques d'enseignement, et ont quitté leurs écoles et collège pour se former.

Merci à tous les membres du LPNC et du GIPSA-Lab où j'ai passé ces trois dernières années. Merci à tous pour les échanges aussi bien scientifiques que personnels toujours intéressants !
Merci tout particulièrement à Coriandre pour son aide technique, Silvain pour les statistiques et Anne-Laure pour le temps passé sur les alignements. Merci à Karine, Marcella et Frederique qui m'ont confié leurs étudiants, me permettant de continuer à enseigner. Merci à Emilie et Sonia pour leurs conseils sur les tracas avec l'Education Nationale et le reste...
Merci également aux membres du LARAC que j'ai eu l'occasion de croiser : à Pascal Bressoux de m'avoir formée aux modèles multiniveaux, à Maryse Bianco pour nos échanges sur la compréhension écrite et à Carole Hanner pour toutes nos discussions à Toronto !

Un grand merci à Omar, Lucrèce, Laura et Brice d'avoir si agréablement partagé mon bureau. Les moments de débats, de rire (et même de ragots) m'ont permis de décrocher et déstresser un peu ! Aprés trois années avec vous, le changement est trés difficile !

Merci aussi à mes amies instits, particulièrement Marie-Ange et Cécile, qui m'ont encouragée et ont fait germer plein d'idées. Merci à Carine qui vient d'obtenir son diplome d'othophoniste, c'était sympa d'échanger sur nos vies d"étudiantes" 15 ans aprés. Merci à la bande de l'EEG, Carine, Arnaud, Aurélie, Thierry, Steph et Nico et à tous les enfants, pour les randos, les soirées, les apéros. Vous avez réussi à me faire décrocher complètement pendant une semaine cet été encore !

Merci à mes parents pour leur soutien, le changement d'air, la garde des enfants et tout le reste !

Un immense merci à mes trois enfants préférés, Martin, Mélisse et Juliette qui m'on soutenus et accompagnés durant ces trois ans. Ils ont prétesté les textes, m'ont acceuillie dans leurs écoles, et laissée expérimenter sur leurs copains... Ils ont accepté que maman parte en conférence, ne soit pas toujours trés disponible pour eux et travaille beaucoup, particulièrement ces derniers temps. Ils ont été super autonomes pendant le confinement ce qui m'a permis d'avancer mon travail malgré tout. Je suis super fière de vous !

Finalement merci à Nico qui m'a toujours suivi dans mes lubies et mes envies de changement, mes hauts et mes bas. Il m'a accompagné, soutenu et supporté. Il a réussi à rester prévenant tout en étant plus que multitaches ces derniers temps et je lui en suis infiniment reconnaissante. Merci pour les cours de LaTeX et la mise en page, et surtout pour tout le reste. Sans toi tout ça n'aurait pas été possible !

Table des matières

Introduction **1**

1 Partie théorique : lecture et prosodie **5**
1.1 Définition, développement et acquisition de la prosodie en lecture . . . 5
Article : A review of reading prosody acquisition and development . . . 8
Prosody and reading . . . 9
Reading prosody development in young readers . . . 15
Reading prosody and comprehension . . . 23
Conclusions . . . 28
1.2 Nouvelles études sur la prosodie en lecture . . . 30
1.2.1 De nouvelles données longitudinales . . . 30
1.2.2 La prosodie en lecture au collège et au lycée . . . 31
Une compétence toujours en développement . . . 31
La confrontation à des types de textes plus nombreux et plus complexes . . 32
1.2.3 Lier prosodie compréhension . . . 33
Évaluation . . . 33
Intérêt pour les publics particuliers . . . 34
1.2.4 Le casse-tête de l'évaluation de la fluence incluant la prosodie . . . 35
1.3 Synthèse des enjeux et questionnements en suspens . . . 36

2 Introduction à la partie expérimentale **39**
2.1 Présentation des études . . . 39
2.1.1 Caractérisation du développement de la prosodie par évaluation subjective . 39
2.1.2 Caractérisation du développement de la prosodie par l'analyse des pauses et de la respiration . . . 40
2.1.3 Caractérisation du développement longitudinal de la prosodie par évaluation automatisée . . . 41
2.2 Précisions méthodologiques . . . 42
2.2.1 Participants . . . 42
Corpus transversal . . . 42
Corpus longitudinal . . . 42
2.2.2 Matériel . . . 44
Textes pour l'évaluation de la prosodie . . . 44
Vitesse de lecture NMCLM . . . 45
Vocabulaire . . . 45
Raisonnement non-verbal . . . 46
Compréhension orale . . . 46
Compréhension écrite . . . 46
2.3 Protocole de passation . . . 48
2.3.1 Passation individuelle . . . 48
2.3.2 Passation collective . . . 49

3 Caractérisation du développement de la prosodie par évaluation subjective **51**
3.1 Étude 1 : Étude transversale du développement de la fluence en lecture incluant la prosodie . . . 51
3.1.1 Présentation de l'échelle . . . 52

Article : Échelle Multi-Dimensionnelle de Fluence : nouvel outil d'évaluation de la fluence en lecture prenant en compte la prosodie, étalonné du CE1 à la 5e . 53
Évaluation de la prosodie . . . 55
Méthodologie . . . 57
Matériel . . . 59
Résultats . . . 61
Discussion . . . 65
Conclusions . . . 67
3.1.2 Différents types d'évaluateurs . . . 68
Méthodologie . . . 68
Résultats . . . 68
Discussion . . . 70
3.1.3 Développement de la prosodie . . . 70
Méthodologie - analyse statistique . . . 71
Résultats . . . 71
Discussion . . . 73
Résultats majeurs Étude 1 . . . 75
3.2 Étude 2 : Étude du lien entre évaluation subjective de la prosodie et compréhension écrite . . . 76
3.2.1 Méthodologie . . . 76
Données expérimentales . . . 76
Analyses statistiques . . . 76
3.2.2 Résultats . . . 78
Fluence . . . 78
Compréhension . . . 79
3.2.3 Discussion . . . 81
Résultats majeurs Étude 2 . . . 83
3.3 Conclusion . . . 84

4 Caractérisation du développement de la prosodie par l'analyse des pauses et de la respiration 85
4.1 Étude 3 : Pauses et respiration du CE1 à l'adulte expert . . . 86
Article : Pausing and Breathing while Reading Aloud . . . 88
Introduction . . . 88
Method . . . 92
Results . . . 95
Discussion . . . 104
Conclusions and perspectives . . . 108
Résultats majeurs Étude 3 . . . 109
4.2 Étude 4 : Marqueurs acoustiques de pauses et perception de la fluence . . . 110
4.2.1 Méthodologie . . . 110
Données . . . 110
Analyses statistiques . . . 111
4.2.2 Résultats . . . 112
Score global de fluence . . . 112
Dimension Expressivité . . . 112
Dimension Phrasé . . . 113
Dimension Décodage . . . 113
Dimension Vitesse . . . 115
4.2.3 Discussion . . . 115
Résultats majeurs Étude 4 . . . 118
4.3 Étude 5 : Marqueur acoustiques des pauses et compréhension . . . 119
4.3.1 Méthodologie . . . 119
Données expérimentales . . . 119
Analyses . . . 120
4.3.2 Résultats . . . 121
Compréhension écrite globale . . . 121
Compréhension en ligne . . . 123

4.3.3 Discussion . . . 123
Résultats majeurs Étude 5 . . . 124
4.4 Conclusion . . . 125

5 Caractérisation du développement longitudinal de la prosodie par une évaluation automatisée 127
5.1 Étude 6 : Développement longitudinal de la prosodie mesurée par évaluation automatisée de la fluence . . . 127
5.1.1 Présentation de l'outil d'évaluation automatique . . . 128
Données audios . . . 129
Evaluation subjective . . . 129
Evaluation objective . . . 129
Validation de la méthode . . . 131
Evaluation automatique . . . 134
Conclusion . . . 134
5.1.2 Développement de la prosodie . . . 135
Méthodologie - analyse statistique . . . 135
Résultats . . . 136
Discussion . . . 139
Résultats majeurs Étude 6 . . . 141
5.2 Étude 7 : Etude longitudinale du développement des dimensions prosodiques de la lecture et compréhension écrite . . . 142
5.2.1 Méthodologie . . . 142
Données . . . 142
Analyses statistiques . . . 143
5.2.2 Résultats . . . 144
Du CE1 au CE2 . . . 144
Du CE2 au CM1 . . . 147
5.2.3 Discussion . . . 151
Résultats majeurs Étude 7 . . . 153
5.3 Conclusion . . . 154

6 Discussion générale 155
6.1 Rappel des principaux résultats . . . 155
6.1.1 Caractérisation du développement de la prosodie par évaluation subjective . 155
6.1.2 Caractérisation du développement de la prosodie par les pauses et la respiration 156
6.1.3 Caractérisation du développement de la prosodie par un nouvel outil objectif et automatisé . . . 156
6.2 Discussion des résultats . . . 157
6.2.1 Evaluer et mesurer la prosodie . . . 157
6.2.2 Caractérisation du développement de la prosodie . . . 159
6.2.3 Prosodie en lecture et compréhension écrite . . . 160
6.3 Limites . . . 161
6.3.1 Dépendance aux textes . . . 161
6.3.2 Préparation des lectures . . . 161
6.3.3 Tests de compréhension . . . 162
6.3.4 Diversité des sujets et des stratégies . . . 163
6.4 Perspectives . . . 164
6.4.1 En recherche . . . 164
Evaluation de la prosodie . . . 164
Cognition et développement . . . 165
Prosodie et compréhension . . . 166
6.4.2 Enseignement de la lecture et application en classe . . . 168

Conclusion 171

Bibliographie 173

Annexes A-1

A Textes utilisés pour les lectures prosodiques **A-1**

B Test de vocabulaire **A-5**

C Tests de compréhension écrite globale **A-11**

D Maze-tests pour la compréhension écrite locale **A-29**

Table des figures

1.1 Proposed time chart showing progressive stages of reading prosody skills development supported by results of the studies reviewed 18

2.1 Organisation matérielle des passations individuelles 48

3.1 Echelle Multi-Dimensionnelle de Fluence (EMDF) 60
3.2 Distribution des scores totaux moyens à l'EMDF pour chaque niveau 63
3.3 Boxplot des scores aux différentes dimensions de l'EMDF en fonction du niveau scolaire 72
3.4 Modéle des liens entre fluence, phrasé et expressivité et la compréhension 79
3.5 Modéle des liens entre compréhension et fluence. Seuls les effets fixes simples significatifs sont représentés ici. Chaque couleur correspond à un modèle 80

4.1 Praat screenshot showing audio signal 94
4.2 Distribution of log duration pauses for each grade level 95
4.3 Boxplots of pause durations corresponding to grade level 97
4.4 Boxplots of pauses frequency according to grade level 99
4.5 Boxplots of breathing pauses from grades 2 to grade 7 and adults 102
4.6 Boxplots of pauses occurring at punctuation marks for readers in grades 2 to 7 and adults 103
4.7 Summary of the results on pausing acquisition in reading development 105
4.8 Interaction nombres de pauses longues-NMCLM dans le score de compréhension écrite en-ligne 123

5.1 Alignement de la courbe mélodique et du coefficient d'allongement des syllabes d'un lecteur à faible score à l'EMDF (en haut) et d'un autre à haut score à l'EMDF (en bas) (courbes rouges) sur un lecteur expert (courbes bleues) dans la lecture du texte 1. 130
5.2 Projection 2D des espaces prosodiques de référence pour le texte 1 avec les lectures de référence (carrés rouges) et celles des enfants (points) et groupes de lecteurs de profils différents 131
5.3 Projection 2D des espaces prosodiques de référence pour le texte 1 avec les lectures de référence (carrés rouges) et celles des enfants (points) et groupes de lecteurs de profils différents 132
5.4 Comparaison des scores moyens donnés par les juges (bleu) et scores prédits (rouge) pour les textes 1 et 2 dans les lectures du corpus transversal, pour chacune des dimensions de l'EMDF 134
5.5 Croissance des scores, évalués par outil automatique, aux quatres dimensions de l'EMDF pour les élèves suivis du CE1 au CM1 137
5.6 Schéma des liens longitudinaux entre compréhension et dimensions prosodiques de la lecture orale du CE1 au CM1 - Chaque couleur correspond à un modèle de l'année n-1 vers n. 151

Liste des tableaux

1.1 Examples of subjective scales used to evaluate reading fluency 14
1.2 Synthesis of the main features of the different studies conducted on reading prosody development 16
1.3 Synthesis of the differences observed between fluent (high reading rate) and less fluent (low reading rate) readers of the same age 19
1.4 Synthesis of the main features of the different studies exploring a link between reading prosody and reading comprehension 25
1.5 Synthesis of the different longitudinal studies exploring a causal relationship between reading prosody and reading comprehension 26

2.1 Description des participants au corpus transversal 43
2.2 Description des participants au corpus longitudinal 44
2.3 Tests écrits réalisés pour chaque corpus pour toutes les années d'enregistrement . . 44

3.1 Caractéristiques des participants en fonction de leur niveau scolaire 58
3.2 Accord inter-évaluateur pour chaque paramètre évalué et pour chaque texte 62
3.3 Moyenne et quantiles des scores de fluence EMDF sur le texte 1 par niveau 62
3.4 Moyenne et quantiles des scores de fluence EMDF sur le texte 2 par niveau 62
3.5 Coefficient de corrélation et corrélation partielle entre les scores EMDF aux textes 1 et 2 63
3.6 Moyenne et quantiles des scores de fluence EMDF sur la moyenne des 2 textes par niveau 64
3.7 Coefficients de corrélation entre score de fluence Evaleo 6-15 et de fluence EMDF . 64
3.8 Accord interévaluateurs (coefficient alpha de Krippendorf) pour chaque type de jury et chaque dimension de l'EMDF 68
3.9 Scores moyens à l'EDMF pour chaque jury et chaque dimension de l'EMDF 69
3.10 Corrélation (coefficient de Spearmann) entre les scores moyens à l'EDMF sur le corpus longitudinal, pour chaque type de jury, deux à deux. 70
3.11 Scores moyens par niveau pour les tests de fluence, compréhensions et vocabulaire. 77
3.12 Comparaison des modèles multiniveaux pour le score global de fluence et les scores aux dimensions de phrasé et d'expressivité 78
3.13 Comparaison des modèles multiniveaux pour le score au test de compréhension globale. 80
3.14 Comparaison des modèles multiniveaux pour le score au test de compréhension en-ligne 81

4.1 Characteristics of the participants according to their grade level 93
4.2 For each grade, parameters of Gaussian mixture models for log duration pauses . . 96
4.3 Comparaison des modèles multiniveaux pour le score global de fluence à l'EMDF 112
4.4 Comparaison des modèles multiniveaux pour le score à la dimension expressivité de l'EMDF 113
4.5 Comparaison des modèles multiniveaux pour le score à la dimension phrasé de l'EMDF 114
4.6 Comparaison des modèles multiniveaux pour le score à la dimension décodage de l'EMDF 115
4.7 Comparaison des modèles multiniveaux pour le score à la dimension vitesse de l'EMDF 116
4.8 Comparaison des modèles multiniveaux pour le score de CE global 121
4.9 Comparaison des modèles multiniveaux pour le score de CE en-ligne 122

5.1 Prédicteurs principaux des scores aux évaluations subjectives et corrélation pour les régressions avec paramètres verbaux uniquement et régressions avec paramètres verbaux et prosodiques. 133
5.2 Scores prédits à l'EMDF pour le texte 1 par l'outil automatique pour chaque année du corpus longitudinal 136
5.3 Modèle de croissance du CE1 au CM1 à la dimension expressivité 136
5.4 Modèle de croissance du CE1 au CM1 à la dimension phrasé 138
5.5 Modèle de croissance du CE1 au CM1 à la dimension décodage 138
5.6 Modèle de croissance du CE1 au CM1 à la dimension vitesse 139
5.7 Scores moyens aux tests de compréhension et scores moyens aus dimensions de la fluence prédits par l'outil automatique pour chaque année du corpus longitudinal . 143
5.8 Modèles multiniveaux pour les scores d'expressivité et de phrasé de l'année 2 . . . 145
5.9 Modèles multiniveaux pour les scores de décodage et de vitesse de l'année 2 146
5.10 Modèles multiniveaux pour les scores de compréhension globale et locale de l'année 2147
5.11 Modèles multiniveaux pour les scores d'expressivité et de phrasé de l'année 3 . . . 148
5.12 Modèles multiniveaux pour les scores de décodage et de vitesse de l'année 3 149
5.13 Modèles multiniveaux pour les scores de compréhension globale et locale de l'année 3150

Liste des acronymes

BP *Breathing Pauses* Pauses respiratoires

CE *Compréhension écrite*

CERGA *Comité d'Ethique de la Recherche Grenoble Alpes*

CO *Compréhension Orale*

CWPM *Correct Word read Per Minute* Nombre de mots correctement lus par minute

DIP *Délai Inspiration-Phonation*, durée comprise entre le début de l'inspiration et le début de la vocalisation du mot suivant la respiration

EMDF *Echelle Multidimensionnelle de Fluence* , échelle d'évaluation subjective de la fluence adaptée pour le français depuis la MDFS

IPD *Inspiration to Phonation Delay* Délai inspiration phonation

LP *Long Pauses* Pauses longues

MDFS *MultiDimensional Fluency Scale*, échelle d'évaluation subjective de la fluence en 4 dimensions proposée par Rasinski et al (Rasinski, 2004; Rasinski *et al.*, 2019)

NMCLM *Nombre de Mots Correctement Lus par Minute*

NMCL *Nombre de Mots Correctement Lus*

NMI *Nombre de Mots Incorrects par minute*

NPIM *Nombre de Pauses Intra-Mots par minute*

NR *Nombre de Répétition par minute*

NV *nombre de Noyaux Vocaliques par minute*

SP *Short Pauses* Pauses courtes

UP *Ungrammatical Pauses* Pauses non grammaticales

UBP *Ungrammatical Breathing Pauses* Pauses respiratoires non-grammaticales

Introduction

"L'homme qui lit à voix haute nous élève à hauteur du livre. Il donne vraiment à lire."

"Etrange disparition que celle de la lecture à voix haute. [...] Plus le droit de se mettre les mots en bouche avant de se les fourrer dans la tête ? Plus d'oreille ? Plus de musique ?" D. Pennac. Comme un roman . 1992.

La lecture à voix haute est trés appréciée des tout-petits. La musique d'un livre lu à voix haute les captive et les berce. Elle est ensuite trés pratiquée par les tout jeunes lecteurs quand ils débutent et tatonnent en décodant. Puis en grandissant, avec le développement de l'automaticité et les nouveaux apprentissages, la lecture silencieuse prend le dessus pour les apprentissages comme pour le plaisir, que ce soit à écouter ou à lire. Quelle est la place de la lecture à voix haute dans l'apprentissage de la lecture ? La lecture est au centre des enseignements à l'école primaire et son acquisition conditionne tous les autres apprentissages. Selon le site Eduscol, site du ministère de l'Éducation Nationale donnant des recommandations aux enseignants (https ://eduscol.education.fr/pid23199/programmes-ressources-et-evaluations.html), en cycle 2, de la grande section au CE1, l'apprentissage de la lecture est centré sur l'apprentissage du code et l'automatisation du décodage. La lecture à voix haute est alors encouragée. Elle vise à une identification des mots plus aisée : "aller vers une lecture fluide et rapide, la seule lecture qui permet de comprendre la signification des mots lus, lorsque par ailleurs on la connaît à l'oral". La lecture à voix haute est alors vue comme une aide à la compréhension. Les compétences en lecture sont évaluées par le nombre de mots correctement lus par minute (NMCLM) et impliquent un respect de la syntaxe et l'emploi d'une expressivité appropriée. En cycle 3, l'enseignement de la lecture est centré sur la compréhension des textes. La lecture à voix haute en tant que compétencce disparait. La fluidité est cependant toujours évaluée par le NMCLM pour l'automaticité et de façon plus qualitative par une échelle à quatre niveaux dont les deux derniers concernent le respect de la syntaxe et l'expressivité. On constate ainsi que la lecture à haute voix doit respecter la syntaxe et être expressive. L'expressivité et le respect de la syntaxe ne sont cependant pas définis clairement et paraissent détachés du reste des compétences. Finalement ces compétences spécifiques à la lecture à voix haute semblent délaissées aprés les premières années d'apprentissages.

Au-delà du plaisir d'écouter et de lire, la lecture à voix haute présente des spécificités et un intérêt certain pour l'acquisition des compétences en lecture. Elle ne se limite pas à un nombre de mot lus correctement par minute, mais comme vu précédememnt met également en jeu des compétences prosodiques : un phrasé et une expressivité appropriés. Lire un texte avec un phrasé et une expressivité appropriés vise avant tout à faciliter la compréhension de l'auditeur par une mise en valeur de la syntaxe et des intentions de l'auteur. Un lecteur expert va utiliser l'expressivité pour rendre sa lecture vivante : faire passer des émotions, appuyer des faits importants, mettre en

valeur ou en retrait certaines phrases... Son utilisation adéquate implique donc une compréhension fine du texte. Toutes ces compétences nécessaires dans la lecture orale ont été mises en lien à travers des définitions récentes de la fluence incluant l'automaticité, la prosodie et la compréhension du texte lu (Rasinski, 2004; Kuhn *et al.*, 2010). Ces mises à jour dans la définition de la fluence visent à mettre au premier plan l'importance de la prosodie dans l'apprentissage de la lecture et à remettre en cause l'utilisation du nombre de mots lus par minute comme unique mesure des compétences en fluence. Celle-ci tend en effet à entretenir la confusion entre vitesse et lecture experte, au détriment de la compréhension de l'auditeur et du lecteur. Cet intérêt récent donne lieu à de plus en plus d'études sur le sujet, mettant en lumière les liens entre prosodie et compréhension et l'intérêt de son intégration dans l'enseignement de la lecture.

Les connaissances sur la prosodie en lecture sont encore limitées, particulièrement chez les jeunes lecteurs. Pourtant, afin de la prendre en compte de façon adéquate dans l'enseignement, il est nécessaire d'avoir une bonne connaissance de son développement et de ses interactions avec les autres compétences en lecture. Les mécanismes cognitifs mis en jeu dans la lecture expressive sont nombreux et leur coordination complexe. Un phrasé approprié nécessite une bonne compréhension de la syntaxe du texte et une planification des pauses. Il est également nécessaire de coordonner le besoin de respirer et cette planification des pauses afin que la respiration ne l'entrave pas. Une lecture expressive nécessite une compréhension fine du texte, des intentions des personnages et de l'auteur, et leur traduction sous forme de variations de mélodie et d'intensité de la voix pendant la lecture. La production d'un phrasé et d'une expressivité appropriés mettent en oeuvre des ressources cognitives importantes et ne peuvent donc se développer qu'une fois ces ressources disponibles. Il est également nécessaire de coordonner la planification de la prosodie avec le décodage et la production de parole. Par ailleurs ces apprentissages commencent à un âge ou les mécanismes respiratoires et de production de la prosodie en parole ne sont pas encore matures. L'étude de la mise en place de ces mécanismes et de leur prérequis permettrait de mieux comprendre le développement de la prosodie en lecture, mais également les mécanismes mis en jeu dans la lecture en général.

L'exploration du développement de la prosodie en lecture, de sa coordination avec les autres compétences, automaticité et compréhension, implique de disposer d'outils de mesure et d'évaluation suffisamment fidèles et sensibles. Or une des spécificités de la prosodie en lecture est la variété des productions possibles. En effet, il n'existe pas une unique façon de lire avec expressivité. Il existe autant de lectures expressives que de lecteurs. L'auditeur va également avoir une perception subjective : une expressivité appropriée pour l'un peut paraitre déplacée, voir non existante pour un autre. Il s'agit donc ici d'avoir une mesure qualitative de la lecture. Si une mesure générale de l'expressivité et du phrasé peut suffire à explorer les liens entre les compétences et à suivre les acquisitions d'un élève, l'étude des mécanismes cognitifs nécessite des mesures plus fines, par exemple de la planification des pauses et de la coordination avec la respiration. Il est possible que cette difficulté à évaluer la prosodie explique que sa prise en compte soit récente.

Au vu des premières connaissances acquises sur le sujet, une meilleure prise en compte de la prosodie en lecture dans l'enseignement de la lecture semble nécessaire. Cela passe avant tout par une meilleure connaissance de son développement et de ses liens avec les autres compétences, ainsi que par la mise en place d'outils de mesure utilisables en classe. Ainsi l'objectif de cette thèse est de caractériser le développement de la prosodie en lecture et de son lien avec la compréhension en proposant différents outils de mesure adaptés à chaque objectif. Pour atteindre ces objectifs, les travaux présentés tout au long de cette thèse se situent dans une approche fortement pluridisciplinaire, alliant la psychologie du développement cognitif et les sciences du langage et de l'ingénieur.

En effet, pour étudier les mécanismes du développement de la prosodie en lecture chez l'enfant et ses liens avec la compréhension (questions relevant du domaine de la psychologie du développement et des apprentissages), il a été nécessaire de développer et d'affiner différentes méthodes et outils de mesure de la prosodie en lecture. Le premier chapitre présentera une revue des études existantes sur le développement de la prosodie en lecture et de son lien avec la compréhension écrite, ainsi que les enjeux actuels dans ce domaine. Le chapitre deux introduira les contributions expérimentales de cette thèse. Il présentera les objectifs des études expérimentales menées et les choix méthodologiques associés. Le chapitre trois sera centré sur l'étude du développement de la prosodie en lecture, et ses liens avec la compréhension, au moyen d'une échelle d'évaluation subjective. Les qualités et les limites de cette échelle, d'usage simple et rapide, seront exposées. Le quatrième chapitre sera consacré aux marqueurs acoustiques de la prosodie et plus précisément à l'étude des pauses et de la respiration. Nous aborderons l'acquisition du phrasé à travers l'anticipation des pauses et la coordination respiratoire. Nous verrons si ces indicateurs peuvent être liés à l'évaluation subjective de la prosodie et à la compréhension écrite. Le cinquième chapitre présente un outil d'évaluation automatisée, se basant sur des paramètres verbaux et prosodiques. L'utilisation de cet outil dans le suivi longitudinal de jeunes lecteurs permet d'explorer la croissance des compétences prosodiques en lecture et l'évolution de leur lien de causalité avec l'automaticité et la compréhension. Finalement, le sixième et dernier chapitre présentera une synthèse des études réalisées, une discussion des résultats obtenus et les perspectives ouvertes par ceux-ci.

Chapitre 1

Partie théorique : lecture et prosodie

La prosodie en lecture est un sujet d'intérêt relativement récent, aussi bien en recherche qu'en contexte scolaire. Le terme même de prosodie en lecture fait l'objet d'interprétations différentes, aussi bien parmi les divers interlocuteurs que dans la littérature. Le lien entre prosodie et fluence mérite également d'être posé clairement avant toute considération sur ce sujet. Ce premier chapitre pose les bases de la définition de la prosodie en lecture et présente une synthèse de la littérature existante. La première partie de ce travail a été publiée dans un article présenté dans le premier paragraphe. Nous compléterons cet article par une synthèse de publications postérieures et d'enjeux non abordés dans l'article mais méritant qu'on s'y arrête dans le cadre de ce travail.

1.1 Définition, développement et acquisition de la prosodie en lecture

Cette première partie est composée de l'article suivant, publié en 2020 dans la revue Reading and Writing : An Interdisciplinary Journal.

Godde, E., Bosse, M. L., & Bailly, G. (2020). A review of reading prosody acquisition and development. Reading and Writing : an Interdisciplinary Journal, 33(2), 399-426. https ://doi.org/10.1007/s11145-019-09968-1

Cette article traite principalement de la définition, l'évaluation, le développement de la prosodie en lecture et sa relation à la compréhension écrite. En parole spontanée, la prosodie (rythme, intonation et intensité de la voix) joue un role essentiel dans la communication du message du locuteur à l'auditeur. Elle a des fonction aussi bien linguistiques (segmentation, mise en relief et hirarchie d'unités...), que paralinguistique (position du locuteur, états mentaux et affectifs...) et non linguistique (poids, âge, taille, sexe...). La prosodie est essentielle pour le développement du langage chez les jeunes enfants : c'est en grande partie à partir des signaux acoustiques que les mots, la syntaxe, la sémantique et la pragmatique du langage ambiant sont appris (Morgan et Demuth, 2014). On s'attend ainsi à ce qu'elle joue également un rôle important en lecture.

Bien qu'on définisse souvent la lecture prosodique comme une lecture "sonnant comme de la parole spontanée", la prosodie en lecture a des caractéristiques légèrement différentes de la prosodie en parole. La prosodie en lecture est parfois incluse dans les déterminants de la fluence au même titre que la précision et l'automaticité. Dans les outils proposés dans la littérature pour évaluer la fluence en prenant en compte la prosodie, les multiples niveaux de structuration de la prosodie sont souvent regroupés en deux composantes : phrasé et expressivité.

Le phrasé recouvre les paramètres rythmiques de la parole : pauses, allongements et intonations, et permet le découpage du texte en unités linguistiques (syntagmes, propositions, phrases, paragraphes...) facilitant la compréhension.

L'expressivité recouvre les variations de mélodie et d'intensité et permet de donner vie au texte, en attirant par exemple l'attention de l'auditeur vers un mot ou un autre ou exprimant l'état d'esprit du locuteur, de l'auteur, voire du personnage.

La variabilité et la subjectivité de la prosodie, et particulièrement de l'expressivité, rendent son appréhension et son évaluation difficile. La mesure de la prosodie, permettant son étude et son évaluation, repose sur deux principales méthodes :

l'analyse objective de paramètres acoustiques (temps et fréquences des pauses, pentes et amplitudes des variations de F0...) et de leur distribution en référence à des lectures adultes.

l'évaluation subjective à l'aide d'échelles permettant à l'évaluateur de noter la maitrise de diverses composantes nécessaires à la communication agréable et facile du contenu du texte.

Ces évaluations, qu'elles soient objectives ou subjectives, restent cependant imparfaites et soumises à interprétation. Les études explorant le lien entre ces deux types de mesures ont mis en lumière la complexité de ces métriques.

Nous avons recensé 13 études s'intéressant au développement de la prosodie en lecture en quatre langues. Le plus grand nombre porte sur l'anglais et les premières années d'apprentissage de la lecture. Ces études utilisent aussi bien l'évaluation subjective qu'acoustique et ne semblent pas mettre en évidence de différence majeure entre les langues étudiées. Leur synthèse amène les conclusions suivantes :

- Le décodage et l'automaticité, deux premiers piliers de la fluence, sont un prérequis nécessaire dans le développement de la prosodie.
- Les placements, durées et fréquences des pauses commencent à se mettre en place dès le CP et leur acquisition se poursuit jusqu'en CE2. Au delà, les réalisations se rapprochent beaucoup des productions adultes.
- L'acquisition des variations mélodiques typiques des phrases interrogatives débute plus tard et continue à se développer au delà de l'éducation primaire.

De rares études s'intéressent à l'expressivité et aux variations d'intensité. Celles-ci semblent apparaitre beaucoup plus tard, en fin de primaire et continuent donc à se développer ultérieurement. Ces études montrent également une grande variabilité au sein d'un même niveau scolaire, en lien notamment avec une grande variabilité dans les compétences de décodage et d'automaticité. L'un des rares modèle de développement, proposé par Miller et Schwanenflugel (2008), souligne l'importance de la prosodie. Le nombre de pauses non-grammaticales et l'amplitude des variations mélodiques sont de bons prédicteurs de la montée en compétence en lecture, en fluence comme en compréhension. Les données analysées ici sont toutefois limitées à peu de langues et peu de niveaux scolaires.

De nombreuses études lient également la prosodie et la compréhension en lecture, directement et indirectement, chez les enfants comme chez les adultes, en lecture orale comme silencieuse. Ce lien

est démontré aussi bien en utilisant des scores subjectifs que des paramètres acoustiques, notamment les paramètres déjà mentionnés, nombre de pauses et variations mélodiques. Ce lien, plutôt faible chez les plus jeunes, se renforce au cours du temps et du développement des compétences en lecture. De plus, chez les enfants et les adultes, la prosodie est accentuée dans la lecture de texte complexe. La question de la causalité est récurrente. La lecture prosodique est-elle une résultante de la compréhension d'un texte ou aide-t-elle à cette compréhension ? Quelques études longitudinales ont été réalisées chez de jeunes apprenants afin de trancher cette question. Les résultats ne sont cependant pas congruents, en partie car les protocoles et les âges des enfants diffèrent. Une synthèse de ces études permet d'émettre un hypothèse sur une possible relation bidirectionnelle entre prosodie et compréhension qui évoluerait avec le développement des compétences en lecture.

Étant donné les liens complexes entre prosodie et compréhension, le manque de données disponibles pour des langues variées et des lecteurs plus expérimentés, et surtout l'implication qu'une meilleure connaissance des liens et du développement pourraient avoir sur l'enseignement de la lecture, il est nécessaire de continuer à explorer ce sujet.

A review of reading prosody acquisition and development

Abstract

The present work reviews the current knowledge of the development of reading prosody, or reading aloud with expression, in young children. Prosody comprises the variables of timing, phrasing, emphasis and intonation that speakers use to convey meaning. We detail the subjective rating scales proposed as a means of assessing performance in young readers and the objective features of each as markers of progress. Finally, we review studies that have explored the intricate relations between automaticity, prosody and comprehension.

Keywords Reading prosody · Reading development · Reading comprehension · Reading fluency · Expressivity

There is a growing interest in the deepening our understanding of reading fluency and developing strategies for improving reading development in young readers. Fluency has long been defined as the ability to read rapidly and accurately, notably through automatic word recognition. Over the past few decades, a second key component of fluency has emerged : reading prosody. We now know that fluency is more than just a question of accuracy and speed but considers the communicative purpose of reading aloud and the role of the listener. Indeed, prosody has been described as "the music of speech" (Wennerstrom, 2001). The length and frequency of pauses, as well as the organization of phonemes duration give rhythm to speech. Variations of the fundamental frequency (F0), phoneme duration and intensity together with other verbal and coverbal cues shape speech rhythm and intonation. The way these cues encode different phonological units and structures varies among languages (Hirst et Di Cristo, 1998), but also among people (Bolinger et Bolinger, 1989). Despite these inter-individual variations, decoding these precious prosodic indications provided by our interlocutors reveals not only the discourse structure but informs the person's emotional, mental and socio-psychological state. The functions of prosody in speech are essential from a communicative point of view (Bolinger et Bolinger, 1989; Wennerstrom, 2001)). Firstly, prosody has an essential linguistic function : intonation and rhythmical structure (e.g., pauses) supplement content-based lexical and morphosyntactic information with redundant and complementary cues that help the listener to parse, disambiguate, understand and remember what is said (Frazier *et al.*, 2006). It also has a paralinguistic function—such as emphasizing a word to focus the listener attention to various linguistic units (words, phrases ...), or adding sarcasm to convey the opposite of what is actually said. Finally, prosody also provides us with information about the speaker—for example, by revealing his or her emotional and physiological state—and how this may be affected by the content. These two last functions of prosody are clearly summarized by Bolinger et Bolinger (1989) (p. 1) :

> Even when it interacts with such highly conventionalized areas such as morphology and syntax, intonation manages to do what it does by continuing what it is, primarily a symptom of how we feel about what we say, or how you feel when you say it.

As prosody is essential for communication, sensitivity to prosody is crucial in speech development (e.g. Schreiber, 1987). Babies rely on prosody to understand speech and learn to speak (de Boysson-Bardies, 1996; Morgan et Demuth, 2014). Developing sensitivity to their native language and in particular to the mother's prosody begins very early in life. The motherese used

with infants eases the acquisition process (Harris, 2013; Nelson *et al.*, 1989). In turn, infants' babbling is also colored by their native language prosody, even before they utter their first sentences (de Boysson-Bardies *et al.*, 1984; Prieto et Esteve-Gibert, 2018). The development of prosody continues until early adolescence (e.g., see Filipe *et al.* 2017, in Portuguese or Wells *et al.* 2004, in English).

Prosody, especially speech rhythm, also plays a role in early reading development. Initially, it thought to be a consequence of the process of developing vocabulary and phonological awareness (Wood *et al.*, 2009). More recently, however, it was shown that prosody on its own can be predictive of variance in word reading at the first grade level in both English (Holliman, 2014; Holliman *et al.*, 2017) and Spanish (Calet *et al.*, 2015a). Furthermore, it was shown that dyslexic readers show, among other impairments, lower prosodic sensitivity than typical readers, both in English (Goswami *et al.*, 2002, 2011) and Spanish (Cuetos *et al.*, 2018).

Just as prosody in speech is essential for understanding, it is natural that prosody when reading aloud, similarly, enhances comprehension of what is being read for both the reader and the listener. Young readers often have the tendency of reading too fast. They believe that the goal of reading is to reach the end of the text as fast as possible, which is not in fact what it means to be a "good reader". In the early 1990's, this observation by Dowhower (1991) would lead prosody to be added to the definition of fluency in reading. An expert reader should in actuality read in the same manner he or she speaks, by using appropriate rhythm and intonation that allow the listener to easily process the speech and understand the content. Adding expression to what is being read is critical for emphasizing, focusing attention and conveying emotions, which effectively are what more recent models of fluency assessment have come to include (e.g. Kuhn *et al.*, 2010; Rasinski, 2004), placing particular emphasis on phrasing and expressivity. The readers likely develop these skills along with other reading skills as they progressively improve their reading fluency. Given the importance of prosody in speech, reading acquisition and overall comprehension, it would seem obvious to understand precisely what the development of reading prosody involves. Unfortunately, the number of studies on prosody as a marker of fluent reading is limited. Prosody does, however, emerge in other studies on reading fluency but is described by different terms. To our knowledge, no previous paper has proposed a comprehensive review of prosodic reading development. Our objective in the present work is to fill that gap by providing an overview of previous investigations on prosodic reading development and its role in young readers' acquisition of reading, and further, to examine how prosodic reading skills and comprehension develop in parallel and are complementary to one another throughout the reading experience.

Prosody and reading

Spontaneous speech prosody versus reading prosody The need to specifically define reading prosody stems from the observation that prosody differs from spontaneous speaking and reading aloud performance. Indeed, while the goal of a fluent reader is to sound like someone talking spontaneously, a closer look at reading prosody reveals a number of differences between these two conditions of speech production (Guaïtella, 1999).

First of all, the distribution of pauses and their durations differ between spontaneous speech and reading. In read speech, pauses are more likely to occur at major boundaries and are shorter and less frequent than in spontaneous expression (Goldman *et al.*, 2009; Grosjean et Collins, 1979;

Hirschberg, 2000; Lalain *et al.*, 2016). Pauses mark a lawful prosodic structure, as strongly related to punctuation (Guaïtella, 1999) as to paragraph structure (Bailly et Gouvernayre, 2012), that do not exist in the planning of spontaneous speech. While articulation rate has also been reported to be higher in reading, results from studies in different languages have differed : Goldman *et al.* (2009) in a study in French and a study by Hirschberg (2000) in English reported higher reading rates, while Beinum (1991) found the opposite to be true in Dutch. Finally, Goldman *et al.* (2009) reported a narrower melodic pattern in spontaneous speech when compared to reading speech in French, and Howell et Kadi-Hanifi (1991) found that stress patterns differed. In English, Cowie *et al.* (2002) posited that this contrast depends on context, as emotional expression is more likely to raise mean pitch height and intensity when spontaneously performed.

There is also significant variability in reading style among readers and reading situations Cowie *et al.* (2002); Dellwo *et al.* (2015); Howell et Kadi-Hanifi (1991). For example, pause placement is particularly reader-dependent. Dellwo *et al.* (2015) looked at listeners' ability to distinguish between read and spontaneous spoken sentences in German. In spite of the relatively few acoustic differences, their 26 listeners were able to accurately distinguish between read and spontaneous spoken sentences. This led the authors to believe that the cues facilitating this are more subtle or nuanced than just pauses, speaking rates and melodic differences. In Japanese, an acoustic comparison between spontaneous speech and read speech shows that spontaneous speech can be characterized by a reduced spectral space (Nakamura *et al.*, 2008).

To conclude, reading prosody differs enough from spontaneous speech prosody to enable listeners to easily distinguish one from the other. This difference would appear acoustic as well as linguistic. The acoustic markers of that difference are, however, not obvious, largely due to of the broad range of different speaking styles between readers.

Reading fluency and prosody It is important to begin by placing prosody in the context of the reading fluency curriculum. Even if the definitions, terms and features used in the literature to refer to reading prosody are diverse, all authors agree that reading prosody plays a critical role in reading fluency. For Kuhn *et al.* (2010, p. 233), "a second critical component of reading fluency is the ability to read with prosody ; that is, with appropriate expression or intonation coupled with phrasing that allows for the maintenance of meaning."

One of the earliest mentions of reading prosody in the context of defining read- ing fluency comes from Dowhower (1991). Reading fluency had previously been equated to reading rate – that is, the number of words correctly read in 1 min – without considering the contribution of prosody, considered harder to assess, even though speed and fluency do not always correlate (Grosjean & Collins, 1979 ; Valencia et al., 2010). According to Dowhower (1991, p. 166), "prosodic reading is the ability to read in expressive rhythmic and melodic patterns". In that respect, she proposed six relevant acoustic features of mature reading prosody : appropriate pausal intrusion, phrase segmentation and length, phrase-final lengthening, terminal intonation contours and stress. Since then, several authors have stressed the importance of incorporating prosody in reading fluency curriculums (Kuhn *et al.*, 2010; Rasinski, 2004; Rasinski *et al.*, 2009; Schwanenflugel *et al.*, 2004; Schwanenflugel et Benjamin, 2012)). Currently, reading prosody is generally considered to be a reading skill, however its definition lacks precision, notably how it relates to reading fluency. Some definitions, such as that proposed by Kuhn *et al.* (2010, p. 240), consider prosody as one essential component of fluency : "Fluency combines accuracy, automaticity, and reading prosody [...]. It is demonstrated during oral reading through the ease of word recognition, appropriate pacing,

phrasing and intonation." By contrast, Cowie *et al.* (2002, p. 49) consider fluency and expressivity as two separate skills : "[...] fluency displays structured-oriented skills, expressiveness displays sense-oriented skills." They suggested integrating phrasing into fluency and putting aside intonation and expressivity, which is to say, separating syntactic from semantic maturing. Importantly, all authors investigating the question have consistently drawn a distinction between three contributions of prosody to reading fluency : phrasing, intonation and expressivity.

Phrasing An essential function of prosody is to chunk the discourse into meaningful units as a means of facilitating the listener's comprehension. The rhythm of speech, specifically the length of syllables combined with the position and duration of pauses, strongly contributes to this chunking process. In fact, boundary accents, pauses and pre-pausal lengthening cue syntactic grouping and structure. This chunking is occasionally referred to as phrasing (Kuhn *et al.*, 2010; Rasinski *et al.*, 2009), rhythmic organization (Cowie *et al.*, 2002) or syntactic prosody (Erekson, 2010).

Three types of pauses exist that have different uses : breath, syntactic and hesitation pauses (Lalain *et al.*, 2016). Breath pauses are necessary for air intake, can be accompanied by audible breathing noises and are used fairly frequently as discourse markers (Bailly et Gouvernayre, 2012). Hesitation pauses are symptomatic of a cog- nitive activity and are largely associated with decoding or planning problems during reading. These frequently translate to ungrammatical pauses. Syntactic pauses aim at highlighting syntactic units to ease sentence parsing and boost comprehension.

Pause position and duration, as well as pause distribution, vary throughout reading acquisition (Lalain *et al.*, 2014). Pause misplacement is often symptomatic of young readers. Hesitation and respiratory pauses are relatively frequent in early reading. Hesitations are mainly due to decoding problems if the reader has not yet acquired reading accuracy and automaticity. Respiratory pauses are also frequently produced by early readers because of their slow speech rate and lower lung volume. Once children have acquired automaticity, they learn to coordinate their breathing and the syntactic parsing of the text. In expert readers, respiratory pauses are often placed at punctuation marks while syntactic pauses further highlight the grammatical structure. Respiratory pauses of skilled readers are generally longer than syntactic pauses and are located at major syntactic boundaries. As reading speed increases, the placement of respiratory pauses may become more anarchic : the need to breathe takes priority over respecting correct grammar. Training children to read fast may then contradict the acquisition of expert reading skills (Grosjean et Collins, 1979; Rasinski, 2006).

In fact, apart from punctuation, there is no prosodic cue in a written text. So punctuation gives only a part of the phrasing (Schreiber, 1980). Most of the phrasing has to be inferred from the syntactic structure of the text, the reader has to manage the syntactic chunking to be able to acquire the correct phrasing, enabling the comprehension of the text.

To summarize, the explicit nature of phrasing that is closely linked to the structure of the text may, therefore, make it the easiest components of reading prosody for young readers to acquire or develop, as long as the reader is able to extract the syntactic structure of the text and perform an online shallow parsing.

Intonation and expressivity Intonation and expressivity are two frequent terms found in the literature when referring to the subjective impact of the melodic variation of the voice. The problem is that these two terms lack precise definition and can be mistakenly interchanged. In the context

of speech prosody studies, intonation refers to the suprasegmental prosody : the variation of pitch linked to linguistic and paralinguistic functions and is invariably specific to the language used (Hirst et Di Cristo, 1998). For example, syllables within a word may be accented in various positions in languages with lexical stress such as English or Spanish. French is considered as a language with no stress (Rossi, 1980), since accents occur at the end of the words. In tonal languages, lexical tones coexist with intonation like waves superimposed on larger swells.

These language particularities are likely to strongly impact reading acquisition and performance. Intonation is mostly linked to punctuation (e.g., it is used to cue declarative versus interrogative sentences). It can also mark asides—cued in the text by commas or brackets—as well as focus (Wennerstrom, 2001). Exclamation marks are also used by authors to induce a variation in the melodic pattern of the sentence. In prosodic phonology, intonation and phrasing are closely linked. In the prosodic structure, the intonation phrase constitutes an intermediate level between the syllable, the word and the intonation unit. The intonation phrase is a stretch (or chunk) of spoken material that has its own intonation pattern. In a nutshell, intonation and phrasing are parts of the linguistic prosody both in speech and reading.

In contrast, expressivity is more subtle. A reader can read with appropriate phrasing and intonation but with no expressivity. His or her reading will be understandable but will lack interest for the listener. Expressivity is in fact linked to paralinguistic prosody (Wennerstrom, 2001). Like intonation, it is encoded through multiparametric variations, pitch, rhythm, intensity as well as timbre, and ultimately it offers implicit information about the text. Authors using the term expressivity generally refer to F0 and intensity amplitude or slope (e.g. Álvarez-Cañizo *et al.*, 2018; Cowie *et al.*, 2002). It can also be cued by punctuation, for example exclamation marks. Indeed, they often trigger expressive reading : the choice of attitude—incredulity, contempt, irony, evidence... (see Golan *et al.* 2006, for an extensive taxonomy of complex emotional displays by face and voice)—depends on the context and the reader's understanding of the situation. Punctuation alone is not sufficient to translate to expressivity in written language so the reader is prompted to infer it from the textual content (Erekson, 2010; Martin, 2011). Erekson (2010) notably insists on the fact that expressive reading implies that the reader has a deep understanding of the text but also infers the emotional state of the character given the context of the story or dialog. This can be explicitly written (e.g., "the boy is mad and he shouts") or not written at all and instead should be inferred from context.

Intonation and expressivity are not always easy to infer from the text alone, and are also highly speaker specific, across language, culture and individuals. This aspect of prosody is the most difficult to teach and acquire by beginners and is also the most challenging to measure in reading assessment.

Assessing reading prosody Reading prosody assessment is an important issue for both teachers and research ers. The evaluation protocol and rated dimensions must be clearly defined to produce reliable results. Rating scales have been used and improved since the 1990's. Generally, this assessment is subjective and relies on expert listeners' perceptual judgments. More recently, acoustic parameters have been used to perform automatic scoring, that predict perceptual judgements quite accurately.

Subjective scales The first rating scales used for measuring prosody were fluency scales that attempted to look further than the reading rate. Such an example is given in Table 1.1.

They were unidimensional scales rating both decoding, phrasing intonation and expressivity (e.g. Pinnell *et al.*, 1995; Zutell et Rasinski, 1991). These scales were mostly developed for the purpose of pupil assessment by teachers. The first three levels focus on the grouping skills. Expressivity appears in the 4th level, assuming that grouping precedes expressivity in the acquisition of reading prosody. Moreover, the term "expressivity" lacks a clear definition, which leaves it up to the assessor's personal interpretation. These unidimensional scales assess several skills in the same level, neglecting the potential variability of development among these skills in children. Consequently, placing a child in a unidimensional scale can be quite difficult.

To improve and facilitate fluency assessment, Zutell et Rasinski (1991) proposed a multidimensional scale, again, conceived to be used by teachers. Rasinski (2004) and Rasinski *et al.* (2009) further updated this multidimensional scale (see Table 1.1) to allow for the separate assessment of four key features : pace, smoothness, phrasing and expression. Each feature is rated on a 4-point scale ranging from poor to correct performance. The problem lays in the fact that pace and smoothness are quite straightforward and easy to evaluate, whereas assessing phrasing and expression can prove more challenging. Indeed, these features tend to differ considerably throughout a given text and both are by nature listener-dependent. A further complication is that assessors may project or be influenced by their own reading strategies. Moreover, without a clear definition of expression, expectations that the assessor may have of a good reader employing appropriate expressivity might also affect the ratings. Actually, when several assessors evaluate the same students, the inter-rater agreement can be satisfactory (i.e., $> .70$) as long as the raters receive appropriate group training (e.g. Moser *et al.*, 2014; Paige *et al.*, 2014; Schwanenflugel *et al.*, 2016). Without this training, the inter-rater agreement tends to be weaker (e.g. Godde, Bailly, Escudero, Bosse, & Gillet-Perret, 2017).

After having investigated the inter-rater consistency of fluency assessment in the literature using the two scales previously presented, Haskins and Aleccia (2014) pointed out a lack of transparency in their use and reliability. They tested their own multidimensional scale to rate fluency. Here again, a 4-point scale was used to assess 6 key prosodic features grouped into two categories : phrasing and expression. Phrasing combines smoothness and punctuation, while expression combines vocal emphasis, inflection, intonation and voice. These features are rated from 1 (no use) to 4 (correct use throughout the entire reading). They asked 83 teachers to rate the prosody of video-recorded children reading the same text. The raters had received no training on the scale. Correlation coefficients between random pairs of raters revealed a range from low (.30) to moderate (.53). Benjamin *et al.* (2013) proposed an acoustically-grounded multidimensional scale detailed in the following paragraph. Using assessments made by 3 raters, all experts in children reading assessment, they still found a moderate inter-rater reliability on exact agreement (even if the reliability on adjacent agreement was better). As a whole, these results attest to a lack of reliability of subjective scales ; they should be used with caution. Moser *et al.* (2014) underline this in their strong recommendation of using at least 2 passages to reach a reliable rating of the pupils being assessed.

To conclude, subjective multidimensional scales primarily dedicated to in-class assessment can be useful for teachers. They may, however, lack the precision or reliability to be used for screening and research purposes.

Objective acoustic markers For researchers, a particular area of interest considers the acoustic markers of child reading prosody. Cowie *et al.* (2002) measured 40 different acoustic markers in the recordings of 8 to 10-year-old readers. They related these acoustic markers to subjective

TABLE 1.1 – Examples of subjective scales used to evaluate reading fluency

Unidimensional fluency scale adapted from Pinnell *et al.* (1995)

Score	Habilities
4	Read with appropriate rate, phrasing and expression
3	Read in small group phrasing, no expression
2	Read in small groups, inappropriate phrasing
1	Read word by word

Multidimensional fluency scale adapted from Rasinski (2004)

Dimension	1	2	3	4
Expression	Non existent	Poor	Mostly correct	Adapted interpretation
Phrasing	Monotonic	Small inappropriate	Run-on and pauses	Appropriate
Smoothness	Frequent pauses	Several rough spots	Occasional breaks	Smooth
Pace	Laborious	Slow	Fast and slow	Conversational

ratings of these recordings. It appears that the acoustic correlates of fluency and expressiveness are those that one would expect by the very definition of the terms. That is to say, fluency is mainly correlated to basic temporal organization : pause duration, pause frequency, syllabic rate and pitch movement frequency. Expressiveness is mainly linked to pitch variation, i.e., pitch movement magnitude and duration and their variation from one sentence to another. However, if fluency markers are primarily linked with temporal organization, they are also linked – to a lesser extent –with expressiveness and vice versa. So even if some acoustic markers seem to reliably characterize fluency and expressiveness, the one-to-one correspondence between objective features and subjective judgment is not that simple. In fact, expressiveness, often linked to pitch variation, also depends on fluency skills. Young readers need minimal fluency skills – such as automaticity – to develop expressiveness skills. These observations on acoustic markers have been confirmed by several other studies (see below in Development section).

Benjamin *et al.* (2013) evaluated the correlation between a new subjective multidimensional scale and acoustic markers of prosody in 2nd and 3rd graders reading assessments. They extracted distributions of several characteristics of the children's prosody : inter-sentential pause lengths, intra-sentential pause ratio, sentence final pitch declination for declarative sentences and pitch contours. A principal component analysis of the objective cues leads to 2 potential distinct prosodic features interpreted as expressive intonation, that is pitch variation variables, and natural pausing, or pause variables. It should be noted that pause variables also showed secondary relations with expressive reading (as in Cowie et al., 2002) for the same reasons previously stated. Features associated with skilled readers (i.e., level 4 on the subjective scale) are larger pitch drops at the end of declarative sentences and larger pitch variation inside the sentences at appropriate places. Features associated with poor readers (i.e., level 1 on the subjective scale) are flat or consist of inappropriate intonation contours and end of sentence pitch variations that are inappropriate with regards to punctuation. Characteristics of intermediate levels (levels 2 and 3 of the subjective scale) are less clear because medium-level readers showed a mix of upper and lower level skills (e.g., they are expressive but not throughout the whole text).

In conclusion, these two studies highlight the difficulty involved in objectively assessing the expressiveness of a reader, especially a medium-level reader, because of the intra- and inter- reader variability. The use of subjective scales, even acoustically grounded, should be used in conjunction with specific training of the raters and should be interpreted with a degree of caution. For research purposes, the use of acoustic parameters to describe reading prosody shows great potential for providing insights into the development of reading prosody. Of additional interest would be to further correlate this objective characterization with subjective ratings with a view to predicting performance from signals without the need of costly listening tests.

Reading prosody development in young readers

Unlike acquisition of decoding and automaticity, reading prosody development has been understudied (see Table 1.2 for an overview). Indeed, we counted only a handful of global studies on reading prosody development including and comparing children of different ages (Álvarez-Cañizo *et al.*, 2018; Miller et Schwanenflugel, 2008; Lopes *et al.*, 2015). Other more specific works (Schwanenflugel *et al.*, 2004, 2015; Benjamin et Schwanenflugel, 2010; Schwanenflugel et Benjamin, 2017; Cowie *et al.*, 2002; Paige *et al.*, 2017) studied the acquisition of given prosodic features (e.g., focus, text complexity, link between fluency and prosody).

TABLE 1.2 – Synthesis of the main features of the different studies conducted on reading prosody development presented in alphabetic order

Paper	Langage	Grade	Measures				Impact of reading fluency
			SR	AA			
				Pause	Pitch	Intensity	
Álvarez-Cañizo *et al.* (2018)	Spanish	3, 5, adults		X	X	X	
Benjamin et Schwanenflugel (2010)	English(US)	2,adults		X	X		X
Cowie *et al.* (2002)	English(UK)	4,5	X	X	X	X	X
Lopes *et al.* (2015)	Portuguese	2,3	X				X
Miller et Schwanenflugel (2008)	English(US)	1 to 2		X	X		
Miller et Schwanenflugel (2006)	English(US)	3, adults		X	X	X	
Paige *et al.* (2017)	English(US)	1,2,3	X				X
Paige *et al.* (2014)	English(US)	9	X				
Ravid et Mashraki (2007)	Hebrew	4		X	X		
Schwanenflugel *et al.* (2004)	English(US)	2, 3, adults		X	X		X
Schwanenflugel *et al.* (2015)	English(US)	3			X	X	X
Schwanenflugel et Benjamin (2017)	English(US)	3			X		
Young et Bowers (1995)	English(Can)	5	X				X

Note : AA acoustic analysis, SR subjective ratings

Main characteristics of the studies We found 13 studies published in English on reading prosody development by conducting an extensive search in Google scholar with the keywords "reading + prosody + development" and by investigating the quotation of Kuhn *et al.* (2010), who introduced a definition of reading prosody.

The main features of these 13 studies are summarized in Table 1.2. Most of the studies are interested in the earlier development of reading prosody during the first years of primary school : 9 out of these 13 studies involve pupils between grades 1 and 3. A common observation of these studies is that pupils first need to acquire decoding and automaticity before progressing further to add prosody to their reading. Only one study (Paige *et al.*, 2014) focused on middle school pupils and the late development of prosody. The studies concerning a single grade often describe the difference of performance between fluent and less fluent readers in each of the acoustic parameters assessed.

Concerning the diversity of acoustic parameters, most of the studies are interested in pause frequency and duration, both intra- and inter-sentential (see Table 1.2). Grammatical pausing was also investigated (Benjamin et Schwanenflugel, 2010; Miller et Schwanenflugel, 2008; Schwanenflugel *et al.*, 2004), most notably in young children. Pitch variation was studied with particular interest in the context of sentence type – in particular, sentence-initial and final pitch movements (Álvarez-Cañizo *et al.*, 2018; Benjamin et Schwanenflugel, 2010; Cowie *et al.*, 2002; Miller et Schwanenflugel, 2008; Schwanenflugel *et al.*, 2004). The alignment of the F0 contour with reference to an adult contour has also been frequently used as a cue for estimating reading development (Benjamin et Schwanenflugel, 2010; Miller et Schwanenflugel, 2006; Ravid et Mashraki, 2007; Schwanenflugel *et al.*, 2004). Intensity is rarely mentioned. With the exception of Cowie's exploration of reading prosody (Cowie *et al.*, 2002), only one study on focus marking (Schwanenflugel *et al.*, 2015) considers intensity variation.

Since all studies focused on use of a single language—mostly English, but also Portuguese or Hebrew—, inter-language comparisons and hypotheses have not been widely developed. Only Álvarez-Cañizo *et al.* (2018) hypothesized about a possible difference between languages. Indeed, we would expect decoding acquisition to be language-dependent. As an example, transparent languages with straightforward grapheme-phoneme correspondences can be decoded rapidly, whereas opaque languages with irregular grapheme-phoneme mapping take longer to learn to decode (Seymour *et al.*, 2003). As such, one hypothesis is that children may acquire prosodic reading earlier in a transparent language (Álvarez-Cañizo *et al.*, 2018). Certain language specificities may also impact the development of reading prosody. For example, in Spanish, interrogative and exclamatory punctuation marks are presented both at the beginning and at the end of sentences, e.g.,"¿Cuándo es la fiesta ?" (When is the party ?). Álvarez-Cañizo *et al.* (2018) hypothesized that, consequently, the pitch contour specific to interrogatives should appear sooner in Spanish than in other more opaque languages.

Major findings of previous studies Taken as a whole, these 13 studies provide us with an overview of the general development of reading prosody studied across elementary grade levels. The evolution of acoustic markers of prosody is summarized in Fig. 1.1. Several of the studies also describe prosodic specificities that can differentiate between poor readers and good readers at the same grade level, as shown in Table 1.3. Our review highlights four important benchmark variables that pave the way for the development of reading prosody : these are acquiring fluency, planning appropriate pauses, choosing the appropriate 'tune' or intonational contour, and finally developing

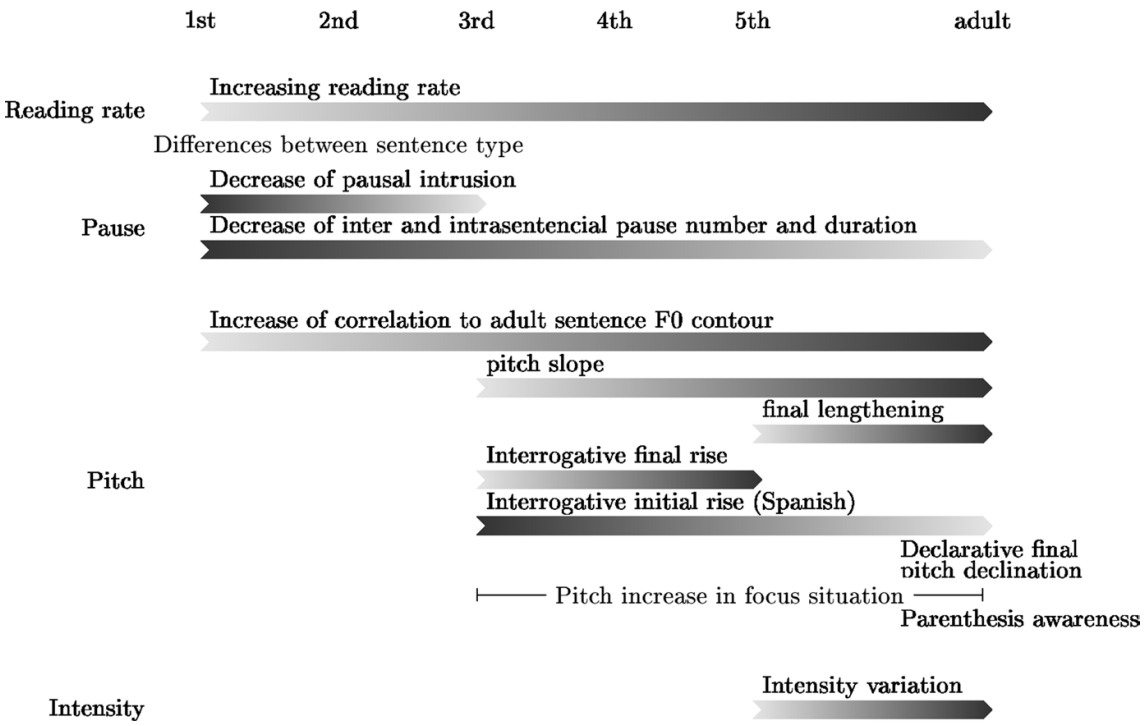


FIGURE 1.1 – Proposed time chart showing progressive stages of reading prosody skills development supported by results of the studies reviewed

expressive language skills. Finally, we cite those works that addressed the large variability of inter- and intra-reader performance and offer insights into the few existing models of reading prosody development.

Fluency The first main conclusion of these studies is the importance of reading fluency (i.e., decoding and automaticity) as a prerequisite for the acquisition of reading prosody. Children become expressive only once they have acquired reading automaticity (e.g. Lopes *et al.*, 2015; Miller et Schwanenflugel, 2008; Paige *et al.*, 2017). Indeed, automaticity reduces the overall cognitive load which effectively frees up attention and cognitive resources—inter alia—that can then be devoted to expressiveness. Consequently, the acoustic markers of appropriate prosody are systematically lower for less fluent readers than more fluent readers (Fig. 1.1).

Pauses Another trend that several studies addressed is the reduction of all kinds of pauses when reading skills increase (Álvarez-Cañizo *et al.*, 2018; Benjamin et Schwanenflugel, 2010; Cowie *et al.*, 2002; Miller et Schwanenflugel, 2008; Ravid et Mashraki, 2007; Schwanenflugel *et al.*, 2004). This trend appears both when comparing children who read at different levels and same-level children with various reading skills. Young children who are poor readers tend to use ungrammatical pauses more frequently and over longer periods. Additionally, the durations of punctuation pauses used by fluent readers frequently decrease. Miller et Schwanenflugel (2008) conjecture that children are less motivated to respect punctuation cues as their reading skills improve. Effectively, adults tend to rely less on punctuation (Chafe, 1988). Ravid et Mashraki (2007) showed that adult-like pausing patterns in Hebrew—that uses a large set of punctuation marks (e.g., sof pasuq, paseq, maqaf...)—appear before intonation, often as early as the 4th grade.

Intonation Intonation as a linguistic function is markedly more difficult to measure than pausing. That said, it has been shown that some features of intonation are acquired quite early in reading acquisition, seen in the fact that initial or final rise in interrogative sentences can be

TABLE 1.3 – Synthesis of the differences observed between fluent (high reading rate) and less fluent (low reading rate) readers of the same age

	Low fluency	**High fluency**
Pauses		
Intra en inter sentential pauses	Long	Short
Complex text	More pauses, more ungrammatical	More pauses, more grammatical
Long text	Increased discontinuities	
Expressivity		
Sentence F0 contour	Flat	Adult-like
F0 rise in interrogative and exclamative sentence	Low	High
Fit to context	Weak	Strong
Complex text	–	More expressivitiy F0 contour, more adult-like
Long text	Weakening	–

present in 3rd graders (Álvarez-Cañizo *et al.*, 2018; Miller et Schwanenflugel, 2006). Nonetheless, intonation at this early stage is not yet adult-like and continues to evolve. On the other hand, other features, such as final lengthening or pitch declination in declaratives, can either appear much later, or can be absent altogether at the 5th grade level (Álvarez-Cañizo *et al.*, 2018). This implies that intonation continues to develop after the fifth grade. Concerning language differences, it was shown that intonation for interrogative and exclamative sentences does not appear earlier in Spanish (Álvarez-Cañizo *et al.*, 2018) than in English (Schwanenflugel *et al.*, 2004), despite the fact that Spanish punctuation marks give strong cues concerning the sentence type.

Finally, the correlation between the F0 contour of children and adults consistently increases from the 1st to 5th grade levels (Álvarez-Cañizo *et al.*, 2018; Benjamin et Schwanenflugel, 2010; Miller et Schwanenflugel, 2006; Ravid et Mashraki, 2007; Schwanenflugel *et al.*, 2004) but is still not very high in 5th grade, which confirms that the acquisition of some aspects of reading prosody continues after the 5th grade.

Expressivity Few expressivity features have been studied in children, which as we have just pointed out, is most likely because it is the most difficult aspect of reading prosody to acquire and to assess. Nevertheless, Schwanenflugel *et al.* (2015) noticed that 3rd graders are able to emphasize words that are meant to be emphasized in various ways : direct quotes, exclamation points or contrastive context. Children do, however, have trouble knowing how to handle parentheses.

Variability The largest body of the studies that focused on studying pupils at a single grade level showed that there is a significant variability in their use of prosody. The expressiveness strategies are different from one child to another. For example, different pitch contours can lead to equally expressive readings, which have led several authors to posit that acoustic markers can only roughly evaluate expressiveness skills (Cowie *et al.*, 2002). This degree of variability is particularly visible between children with different reading skills (see Table 1.3) (Cowie *et al.*, 2002; Lopes *et al.*, 2015; Schwanenflugel *et al.*, 2004). The differences observed between skilled and less skilled readers are the same as those that differentiate children and adults (Binder *et al.*, 2013). Moreover, the degree of variability between fluent and less fluent readers also depends on text complexity and length (e.g. Benjamin et Schwanenflugel, 2010). When a given text is either more complex or longer, variability between fluent and less fluent readers will tends to be higher, notably in terms of expressivity and ungrammatical pausing (Benjamin et Schwanenflugel, 2010).

Less skilled readers make more pauses, appropriate or not. They rely more on punctuation. Regarding pitch, their F0 contour tends to be flatter or more monotone. They also provide poor cues of modalities, for example, by omitting declination lines or falling/rising boundary tones for declarative/interrogative sentences. Cowie *et al.* (2002) pointed out that pitch variations progressively become more important as the readers develop their fluency. Less fluent readers typically exhibit poor pitch variation between sentences.

As slow and poor readers, children with dyslexia or language impairment also present numerous long and ungrammatical pauses compared to control children, their reading rate is slower and their melodic contour is flatter than those of control children (Suárez-Coalla *et al.*, 2016; Jordán *et al.*, 2019; Lalain *et al.*, 2012). These studies support the impact of decoding and automaticity issues on fluency. The different authors point out the use of cognitive resources on decoding at the expense of anticipation, comprehension and prosodic planning.

Developmental model Two studies explored the links between different mark- ers of prosody in order to describe potential dependencies. In a longitudinal study, Miller et Schwanenflugel (2008)

examined the relationship between pauses (pausal intrusion rate) and pitch (F0 contour match to an adult reference), by monitoring both in 1st and 2nd graders. Unsurprisingly, they found that pauses and pitch recorded in 1st grade pupils are respectively correlated to pauses and pitch measured one year later. More importantly, the pausal intrusion rate in first grade is also related to the F0 match in second grade. Indeed, the children having fewer pausal intrusions in first grade present a more adult-like F0 contour in the second grade. This result suggests that a decrease of pausal intrusion is a precursor of improved intonation.

In the same study, Miller et Schwanenflugel (2008) also tested the contribution of pausal intrusion, F0 match and word reading skills recorded in 1st and 2nd grade pupils, as predictors of 3rd grade level reading fluency. As expected, word reading skills observed in at the 1st and 2nd grade levels are related to reading fluency skills seen in the 3rd grade. However, both F0 match in 1st and in 2nd grade pupils also proved to be important predictors of reading fluency in the 3rd grade once early reading skills were taken into account. Thus, it would appear that the early development of intonation contour plays a crucial role in the development of later stage fluent reading. Schwanenflugel *et al.* (2015) also examined the relationship between reading rate, accuracy and different prosodic features in 3rd grade readers. This study confirmed that throughout the development stage, reading rate and accuracy are correlated to reading prosody markers such as intra-sentential pausing or pitch change.

Discussion The present review of studies that have focused on reading prosody in children provides an overview of the main rules governing reading prosody acquisition, such as (1) that fluency – mostly automaticity – is necessary to begin to enhance prosody, (2) that an early decrease of pausal intrusions allows for later enhancement of intonation, and (3) that the early ability to read with appropriate intonation and expressivity can fairly accurately predict later stage reading skills.

There are, however limitations to the developmental synthesis proposed here. Importantly, there is a lack of data from various sources because several of the 13 studies reviewed here were conducted by the same teams and were part of the same project. There is also a noticeable gap in the literature of investigations focused exclusively on reading prosody as an isolated skill ; it has more commonly been addressed as a part of others skills such as reading fluency or comprehension, and has largely been evaluated using only subjective scales.

The lack of available data also stems from the fact that all studies except one focus on primary school pupils, whereas there is data that would suggest that prosody development further continues as young readers enter secondary school. For example, sentence-final boundary tones have been shown to appear after the end of the primary school. In secondary school, fluency in general – prosody in particular – is only very rarely addressed. A series of studies conducted by Paige *et al.* (2012) and Rasinski *et al.* (2005) showed that reading skills have a positive impact on pupils' literacy achievement, particularly in the case of struggling readers. This alone suggests that it would be worth exploring the acoustic markers of prosody produced by secondary school pupils. Moreover the differences in prosodic reading observed in young readers with different fluency skills can also be observed in high-school and young adults (Binder *et al.*, 2013; Paige *et al.*, 2014), and prosody is also related to comprehension in adults (see section Prosody and Comprehension). However, after primary school, reading is barely taught, it becomes only a tool to acquire new skills in many fields, leaving struggling reader with general difficulties. Exploring prosody acquisition after

primary school could lead not only to a better knowledge of its late development, but could also help designing intervention to help older struggling readers.

A comparison between English and Spanish studies suggests that language transparency and punctuation style have little impact on prosody acquisition. It would be interesting to test this assumption further by exploring prosody acquisition across a wider range of languages. Another potentially important line of exploration might be to look specifically at the impact or role of curriculum on reading prosody development. Prosody has been used in the US for much longer than in Europe as an integral part of fluency assessment protocols (e.g., National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.), 2000), and specific training is made available for that purpose. Conversely, fluency in France is still measured largely by gauging reading rate, which effectively confines assessment to accuracy and automaticity. The result is that very little training is offered in prosody itself. It could potentially be interesting to develop reading prosody training programs in different languages and to compare reading prosody development in the same language, notably in children with different curriculums.

This review reveals several important missing points concerning reading prosody development, which certainly reserve further investigation. One of them is the role of syntactic awareness, which could also be an important prerequisite of prosody development. Fluency, word reading, accuracy are often used as predictors of prosody development. However, to produce appropriate phrasing, pupils also need to be able to parse the text into meaningful units (Young et Bowers, 1995). Another interesting point that deserves more attention is the coordination of breathing and reading (Bailly et Gouvernayre, 2012; Bailly *et al.*, 2013; Grosjean et Collins, 1979; Lalain *et al.*, 2014). As mentioned above, breathing is one indispensable motivation for pausing. When poor readers increase their reading speed, breathing tends to become haphazard and induces ungrammatical pausing, disturbing reading prosody. It could also be interesting to investigate when children manage to produce relevant or appropriate breath pauses, that is, when they learn to coordinate breathing and reading.

Both syntactic awareness and breathing coordination necessitate an online analysis of the text and its syntax, as well as anticipation and planning. According to several authors, the reader needs to first acquire decoding and automaticity in order to free cognitive resources which could be allocated to analysis and anticipation. The problems encountered by children with language impairment and dyslexia could come from this (e.g. Suárez-Coalla *et al.*, 2016; Jordán *et al.*, 2019). The anticipation issue has been raised, inter alia, by Schreiber (1980). We didn't found any studies focusing on the anticipation issues and their link with decoding and automaticity skills, with the reading rate or with the age of the reader. Is anticipation and online analysis only a matter of decoding skills or also a matter of cognitive maturity ? This question leads back to the lack of studies in secondary grades. As cognitive maturity has an impact on prosody development, and as the brain is not completely mature at the end of primary grades, studies with elder pupils could bring a lot of additional information on prosody development.

Finally, these studies have well characterized the relation between fluency and prosody with different levels of text complexity. The fact that text complexity enhances prosody in good readers circles back naturally to the role of prosody in reading comprehension. We further explore this issue in the next section.

Reading prosody and comprehension

According to Schreiber (1991) and Morgan et Demuth (2014), children rely on the regular prosodic patterns of speech to perform a shallow syntactic parsing of utterances. When it comes to reading, they have to reverse this process and rely on syntactic parsing and online comprehension to produce adequate prosody and fluent reading. Meisinger *et al.* (2009) warn about the downsides of intensive training to increase reading rate. It often improves at the detriment of other reading skills, like prosody and comprehension. It creates what they term "word callers" who can read fluently but with little or no understanding of the text being read. In reality, we expect to see a strong relation between syntax, prosody and reading comprehension given that syntactic grouping is essential to understanding utterances, and a fortiori, the text itself. This relation has been studied in several languages, in both adults and children. The correlation between reading prosody and reading comprehension has been shown to be significant in most studies. However, in the context of development, this relation raises the question of the causal link between prosody and comprehension. Does prosody enable children to understand the text ? Or does understanding the text enable them to adapt appropriate prosodic parameters while reading aloud ?

Prosody and comprehension in adult readers As mentioned earlier in section "Development", prosody is a central part of speech comprehension. Frazier *et al.* (2006) underline the importance of prosodic phrasing in speech comprehension. Two studies, conducted respectively with French and German adults, extend that observation to both oral and silent reading. Dodane et Brunellière (2006) and Kentner et Vasishth (2016) showed that adults covertly recreate prosody when reading silently when confronted with complex or ambiguous sentences. The word reading time in silent reading is correlated to the word lengthening in aloud reading (Dodane et Brunellière, 2006). Kentner et Vasishth (2016) results reveal a strong interaction between text disambiguation and prosodic cues – that is, local sentence rhythm and global context comprehension – in both silent and aloud reading. Both teams hypothesized then that readers covertly recreate in silent reading the prosodic accentuation of aloud reading needed to process complex sentences.

Another important cue of prosody for comprehension is the relationship between syntactic structure, pausing and comprehension. Koriat *et al.* (2002) proposed that prosody acts as a tool for early syntactic extraction. Their study, conducted with Hebrew-speaking students, reveals that the extraction of prosodic structure precedes the analysis of meaning. In this study, reading prosody is linked to the syntactic structure but remains independent from semantic coherence. Along these same lines, Binder *et al.* (2013) showed that skilled readers do not use pauses as frequently as less-skilled readers. Low-skilled adult readers tend to make more pauses, especially at commas where they always produce long pauses, whereas skilled readers will not pause when commas are used, for example, to separate a list of adjectives, but will pause when a comma separates two clauses. This study showed that comprehension and pausing pattern, number, placement and duration are linked.

Taken together, these results indicate that prosody, particularly pause placement and phrasing, is essential to expert reading comprehension.

Link between prosody and comprehension in young readers During reading acquisition, several literacy skills develop simultaneously : accuracy, automaticity, reading rate, prosodic skills and comprehension. As presented in section "Development", these skills are interdependent, but

what is the exact role of comprehension in the acquisition of literacy skills ? Comprehension is often presented as the ultimate goal of reading acquisition. That said, establishing a link between comprehension and the development of other literacy skills is not that straightforward. While a number of previous studies have questioned the relation between fluency and comprehension, it is only more recently that increased attention has been given to the relation between prosody and comprehension.

The first hints in favor of a link between prosody and comprehension in child reading come from the comparison between aloud and silent reading. Several studies report that, in young and poor readers, reading comprehension is better in oral reading than in silent reading (e.g. Dickens et Meisinger, 2016; Prior *et al.*, 2011), while comprehension scores in both conditions are the same for middle-school and skilled readers. Paige *et al.* (2014) reported an impact of prosody on silent reading comprehension in 9th graders. They suggest that prosody serves as a mediator between automaticity and comprehension. They hypothesize that oral reading benefits young readers because prosody helps them to understand the content. Nevertheless, the results may be biased by the fact that the reading rate is significantly faster in silent reading compared to oral reading (e.g. Prior *et al.*, 2011), suggesting that comparing oral and silent reading may be irrelevant for studying how prosody and comprehension relate to each other.

Another hint in favor of a link between prosody and comprehension is the effect of text complexity on reading prosody. Benjamin et Schwanenflugel (2010) studied the impact of text complexity on reading prosody in 3rd graders. It appears that children tend to accentuate prosody while reading complex texts by using both pausing and pitch variation. This effect is particularly salient with good readers. Miller et Schwanenflugel (2006) observed the same trend in 3rd graders who accentuated prosody in the case of complex sentences. Young et Bowers (1995) also showed that phrasal knowledge – measured by the ability to parse a text with meaningful boundaries between phrases or clauses – significantly explains reading fluency and prosody, in particular with difficult text and long sentences. These three studies with children stressed the hypothesis that prosody supports or boosts comprehension.

To conduct this review, we looked specifically for studies conducted to explore reading prosody and comprehension in elementary children in several languages. A search with “reading + prosody + comprehension” in Google Scholar and the investigation of the quotations of Rasinski (2004), – credited with developing the Multidimensional Fluency Scale that is today widely used to study prosody-comprehension links – enabled us to find the studies linking reading fluency and comprehension. We included in our review 11 studies that specifically investigated the reading prosody-reading comprehension link (see Table 1.4), and 5 longitudinal studies (see Table 1.5).

In the studies involving early readers (Lopes *et al.*, 2015; Schwanenflugel *et al.*, 2004), the correlation between reading prosody and reading comprehension happened to be weak. This suggests that for pupils at the 1st and 2nd grade levels, comprehension is mainly related to decoding speed and accuracy. At these grade levels, a low reading rate has a strong impact on comprehension. However, Paige *et al.* (2017) identified a mediating role of prosody in the relationship between automaticity and reading comprehension in early reading acquisition. One hypothesis about reading comprehension of young readers is that they understand the text by listening to themselves (Kuhn *et al.*, 2010; Schreiber, 1991). Based on this same hypothesis, several studies screened for pauses, in particular the proportion of inappropriate pauses and their relation with reading comprehension. For example, Arcand *et al.* (2014) highlight the impact of inappropriate pauses in reading comprehension in 3rd graders independently from reading rate and accuracy. They confirm the

Table 1.4 – Synthesis of the main features of the different studies exploring a link between reading prosody and reading comprehension presented in alphabetic order

Paper	Langage	Grade	Measures		
			SR	AA	
				Pause	Pitch
Álvarez-Cañizo *et al.* (2015)	Spanish	3,6		X	X
Arcand *et al.* (2014)	French	3	X		
Calet *et al.* (2015b)	Spanish	2, 4	X		
Lopes *et al.* (2015)	Protuguese	2 to 3	X		
Miller et Schwanenflugel (2006)	English (US)	3, adults		X	X
Paige *et al.* (2014)	English (US)	9	X		
Paige *et al.* (2017)	English (US)	1, 2, 3	X		
Ravid et Mashraki (2007)	Hebrew	4		X	X
Schwanenflugel *et al.* (2004)	English (US)	2, 3		X	X
Veenendaal *et al.* (2014)	Dutch	4	X		
Yildirim *et al.* (2018)	Turkish	4,5,6,7,8	X		
Yildiz *et al.* (2014)	Turkish	5	X		

Note : AA acoustic analysis, SR subjective ratings
The subjective ratings used the multidimensional fluency scale (Rasinski, 2004)

importance of appropriate pausing for reading comprehension. It should be noted that, even if less investigated with young readers, pitch variation also appear to be linked to comprehension (Álvarez-Cañizo *et al.*, 2015; Schwanenflugel *et al.*, 2004).

After the 3rd grade, the direct correlation between reading prosody and reading comprehension seems to strengthen. This trend was found in various languages : Dutch (Veenendaal *et al.*, 2014), Spanish (Álvarez-Cañizo *et al.*, 2015; Calet *et al.*, 2015b), and Turkish (Yildirim *et al.*, 2018). A strong correlation between inappropriate pausing and comprehension was also confirmed with older students and in several languages : with Spanish 3rd and 6th graders (Álvarez-Cañizo *et al.*, 2015), with 4th graders in Dutch (Veenendaal *et al.*, 2014), as well as in Hebrew (Ravid et Mashraki, 2007). The correlation between pitch and comprehension is also present in older students (Álvarez-Cañizo *et al.*, 2015; Calet *et al.*, 2015b). Studies monitoring middle school pupils—ranging from the 4th to 8th grade—observed that prosody predicts a larger part of reading comprehension as the grade level advances (Yildiz *et al.*, 2014; Yildirim *et al.*, 2018).

To summarize, the relation between prosody and reading comprehension seems weak for young readers, probably because of the importance of reading rate and decoding skills in comprehension at this age. As children acquire better reading skills, however, this relation appears to strengthen.

Prosody and comprehension : causal links ? As we outlined in the previous section, the link between reading prosody and comprehension is observed in many languages and is strengthened by increased reading performance. However, the causal relationship between reading prosody and comprehension in development, if there is one, is not clear in the studies previously mentioned.

TABLE 1.5 – Synthesis of the different longitudinal studies exploring a causal relationship between reading prosody and reading comprehension, presentede in alphabetic order

Author	Language	Grade	Assessment	Measures	Causal links		
					No	Uni directionnal	Bi directionnal
Fernandes *et al.* (2018)	Portuguese	2 to 3, 4 to 5	3 times	MDFS with 6 levels	4 to 5		2 to 3
Klauda et Guthrie (2008)	English (US)	5	2 in 3 months	Subjective scale 5 dimensions			X
Lai *et al.* (2014)	English (US)	2	3 times	NAEP scale	X		
Lopes *et al.* (2015)	Portuguese	2 to 3	4 times	Subjective assessment	X		
Veenendaal *et al.* (2016a,b)	Dutch	4 to 6	3 times	MFDS		Changing with grades	

Note : prosody was measured with subjective ratings using MDFS for multidimensional fluency scale ((Rasi) and NAEP scale (National Reading Panel (U.S.) and National Institute of Child Health and Human Development (U.S.), 2000))

One question is the direction of this causal relationship. Some data suggest that children rely on prosody to boost their comprehension, particularly when confronted with complex texts. The use of emphasis when reading complicated texts or ambiguous sentences has been observed in both adults (e.g. Binder *et al.*, 2013; Kentner et Vasishth, 2016; Koriat *et al.*, 2002) and children (e.g. Benjamin et Schwanenflugel, 2010; Frazier *et al.*, 2006; Miller et Schwanenflugel, 2006; Schimmel et Ness, 2017). As observed by Schreiber (1991), children still rely on prosodic patterns of speech to infer syntactic organization of sentences when they are beginning to learn to read. Their understanding tends to improve when they read aloud (Frazier *et al.*, 2006). One interpretation of these studies is that good readers rely on prosody to understand the text and, as such, will have a tendency to over-emphasize reading difficult passages that are more complex. These observations seem to favor a causal link from prosody to comprehension. The fact that this effect is particularly noticeable with highly skilled readers may suggest, however, that this tendency depends on level of reading skills involved and, consequently, is reflective of the stage of reading development (Calet *et al.*, 2015b).

Longitudinal studies are known to be one of the most reliable methods for exam- ining causal relationships during reading development. According to our research, only five longitudinal studies have aimed specifically at exploring the bidirectional links between reading prosody and reading comprehension (see Table 1.5).

The results of these studies are nearly inconclusive : 2 studies out of 5 did not observe any link between prosody and comprehension. Lopes *et al.* (2015) screened prosody and comprehension in 98 children four times at the 2nd and 3rd grade levels. The correlation between the ratings of these two dimensions was very low. One shortcoming of that study, mentioned by the authors, was the short term follow-up and the repetitive use of the same text at relatively short intervals of time. Lai *et al.* (2014) conducted a longitudinal one-year study with 2nd graders and didn't find any relation between reading prosody and comprehension. Considering that comprehension is mostly linked to fluency in early readers, a possible explanation is that the second graders were not fluent enough to observe an impact of prosody on comprehension at this age.

Nevertheless, a bidirectional link was observed in the other 3 remaining studies. Klauda et Guthrie (2008) were the first to explore the possibility that bidirectional relations exist between fluency—prosody in particular—and comprehension. They indeed found an influence of prosody on comprehension and vice versa in their study observing 5th graders at 3 month intervals. This study has paved the way for longer term studies. Veenendaal *et al.* (2016a) and Veenendaal *et al.* (2016b) studied pupils from the 4th to 6th grade in Dutch and proposed a bidirectional model that fits well with their data. In this model, comprehension at the 4th grade level is positively correlated to prosody in the 5th grade. Then the link is reversed and the prosody at the 5th grade is positively and significantly correlated to comprehension at the 6th grade. They also noticed that the different components of fluency—reading rate, phrasing, intonation and expressivity—have changing correlations to comprehension over time and as reading skills progress. This bidirectional link was also observed in Portuguese between the 2nd and 3rd grades (Fernandes *et al.*, 2018). They also found a predictive effect of prosody related to effortless comprehension (text vs. word reading) but found that evidence of a bidirectional relation between the 4th and 5th grades no longer existed. According to these authors, the relation depends on grade level and may be strongly influenced by the orthographic transparency of the language. Indeed, Portuguese is characterized by an intermediate level of orthographic depth, compared to Dutch which is transparent, which subsequently explains a bidirectional link for Dutch (Veenendaal *et al.*, 2016a) but not for their Portuguese data.

Discussion Several studies have conclusively established that a relationship exists between prosody and comprehension. Once automaticity is in place, this link strengthens. The causal relation, however, is less clear. The link may be bidirectional and may change over time with the development of more advanced reading skills. Unfortunately, that fact that so few studies have focused on the possible causal relations between prosody and comprehension prevents us from gaining a clear picture, notably because the few that have focused on this specific dynamic were unable to produce converging results, for a number of reasons. Age difference between cohorts, differences in language transparency, and methodological variations are among the many possible causes explaining this divergence. We could add other possible reasons like difference of grammatical complexity between languages or difference of reading curricula between countries. More longitudinal studies on prosody and comprehension development would potentially shed important light on this question. Gaining deeper knowledge of the direction of this link could translate into developing new interventions to support readers with poor comprehension.

Syntactic awareness was measured only by Veenendaal *et al.* (2016a) and Klauda and Guthrie (2008), as one of the significant predictors of comprehension. Any link between syntactic awareness and oral reading prosody has not yet been directly studied. Given the use of syntactic awareness to correctly parse discourse, while also essential for appropriate phrasing, it could be worth investigating the link between syntactic awareness, phrasing and comprehension.

Conclusions

Oral reading prosody is a recent area of interest in the reading curriculum. It can be assessed either with subjective scales or by screening acoustic parameters. Subjective scales provide researchers and teachers with a rapid and easy way to assess reading prosody, but these are mostly designed for classroom use and their reliability depends on the availability of multiple trained assessors. To study prosodic parameters in greater detail, as well as the development of the language-specific spatiotemporal patterns, an analysis framework will need to be established that takes into account current prosodic models and theories of intonation.

The development of reading prosody begins with the development of the other reading skills. The studies reported here in different languages offer a timeline that roughly outlines prosody development. Phrasing is the first skill to develop with the reduction of number and lengths of pauses. Once decoding and automaticity are acquired, young readers can focus on the appropriate placement of the pauses, using appropriate intonation and adding expressiveness to their readings. The development begins in early years of reading acquisition and continues until adulthood. Some prosodic features are in fact not yet managed at the end of primary school. More studies with older pupils and in various languages could be useful to provide a wider and more complete view of reading prosody development.

Prosody is linked to comprehension, both in speech and reading. The direction of the link between reading prosody and reading comprehension has not been sufficiently studied to provide a clear picture of perquisites, if any. Nevertheless, this bidirectional link seems to evolve with fluency skills and grade level. A plausible explanation is that the pupils need to understand the text to adjust their prosody to its content in the early stages of reading acquisition. As the reading skills increase, pupils turn back to relying on prosody to enhance their comprehension of the text. As for development, more studies would be useful to understand the precise nature of the relation

between prosody and comprehension. They should also examine if particularities of the language – orthographic, linguistic as well as phonological – have an impact on the prosody-comprehension link.

Future research on reading prosody development and the possible comprehension-prosody link would be useful towards designing new methods of intervention (e.g., see Rasinski et Smith 2018). Indeed, if the bidirectional link is confirmed, improving reading prosody could lead to an implicit improvement of comprehension. Training reading prosody could be a way to help both fluent readers and poor comprehenders. Prosody teaching raises two questions : when and how. In many curriculums, reading instruction focuses firstly on decoding and automaticity and then possibly prosody is taught. Rasinski (2006) and Rasinski (2010) pointed out the issues of teaching decoding and automaticity apart from prosody. This separation tends to create fast readers at the expense of comprehension of both listener and reader. As reading is the combination of decoding, automaticity, prosody and comprehension, Rasinski suggests that every action should be taught together. Actually, this question closely depends on the main factors that influence prosody development. Is it more linked to cognitive maturity, automaticity, comprehension or a combination of several factors ? A more precise development scheme of reading prosody would potentially help determine the best time to work with young readers on reading prosody in order to have a maximal impact on comprehension.

The second issue about prosody training is the intervention design. The first important point is the choice of the texts used. Indeed to improve prosody the text read should lead to prosody, like poetry or theater. According to Young et Rasinski (2018), repeated reading for performance, like Readers Theaters, give motivation to children to improve their prosody. In fact repeated reading focused on prosody, and not on rate, helps children to extract the syntax and meaning of the text and so improve prosody. Another promising intervention is modeling. Listening to an expert expressive reader gives the child guidance to what expressive reading sounds like. The modeling could also be important in pre-reading years, especially when teaching prosody and decoding simultaneously. It is easier for a child to produce prosodic reading, when he/she has a picture of what is prosodic reading. To conclude, prosody in reading has been shown to be an important part of reading skills, just as decoding, automaticity and comprehension. Despite this, its inclusion in the curriculum is not optimized. Future research on its development, its place in the curriculum and the different way to enhance prosodic skills of pupils, young readers and more skilled ones, needs to be conducted.

Funding Funding was provided by Caisse des dépots et consignation (E-FRAN PIA program).

1.2 Nouvelles études sur la prosodie en lecture

Lors de la 69e conférence annuelle de la LRA (Literacy Research Association) en 2019, Rasinski et Paige rappellent combien la fluence, dans sa globalité, est encore peu prise en compte dans l'enseignement de la lecture, délaissée au profit de la correspondance graphème-phonème et de la compréhension. De plus, quand la fluence est abordée, elle est souvent limitée aux seules mesures de vitesse de décodage et de précision, omettant complètement la prosodie, sujet soulevant peu d'intérêt, car, selon certains, inutile à la lecture silencieuse.

Depuis la soumission de la revue présentée ci-dessus, de nouvelles études sont venues renforcer les observations réalisées. L'intérêt pour ce domaine croit et l'enjeu de la prise en compte de la prosodie en lecture dans l'évaluation des compétences et de son enseignement est de plus en plus présent dans la littérature. D'une part, de nouvelles études développementales complètent et renforcent les données existantes. D'autre part, le lien entre prosodie et compréhension s'affine avec l'intérêt pour de nouvelles populations notamment les élèves en difficultés, collégiens, élèves présentant des troubles du langage. Cet intérêt permet d'ouvrir de nouveaux champs d'étude et d'apporter de nouvelles informations dans la construction d'un modèle plus général de développement des compétences en lecture. Finalement, la multiplication des études sur la prosodie en lecture amène également les chercheurs à se questionner sur son évaluation, cherchant à répondre aux questions soulevées dans l'article précédent.

1.2.1 De nouvelles données longitudinales

Deux études longitudinales, publiées trés récemment, apportent un éclairage supplémentaire à la problématique du développement de la prosodie en lecture. Ces deux études portent sur des enfants du CP au CE2.

Dans une première étude portant sur des enfants espagnols, Álvarez-Cañizo *et al.* (2019) ont réalisé une mesure par an et suivi le développement des paramètres acoustiques traditionnellement associés à la prosodie dans la littérature. Cette étude peut être rapprochée de celle réalisée par Miller et Schwanenflugel (2008) en anglais, présentée dans l'article précédent. Álvarez-Cañizo *et al.* (2019) ont pu observer les mêmes tendances en espagnol que celles observées en anglais. Avec l'augmentation du niveau scolaire, on observe une diminution du nombre et de la durée des pauses, ainsi que du nombre d'erreurs de positionnement. Le nombre de pauses non grammaticales diminue uniquement du CP au CE1. Concernant les variations mélodiques, il y a peu d'évolution dans le contour intonatif des phrases déclaratives, mais celui des phrases interrogatives n'apparait qu'en CE1. Les auteurs concluent donc à un développement similaire en anglais et en espagnol avec une étape importante entre le CP et CE1, et ce malgré la différence de transparence des deux langues. Cette étude vient donc renforcer les constatations réalisées dans l'article précédent. Bien qu'on aurait pu envisager un développement plus précoce de la prosodie en lecture dans les langues plus transparentes (étant donné que le décodage est acquis plus tôt), cela ne semble pas être le cas.

La deuxième étude présente une approche différente de celles qui ont servi de base au travail de synthèse mené jusqu'ici. Dans cette étude, des enfants américains ont été suivis deux fois par an du CP au CE2, six mesures ont été réalisées (Young-Suk *et al.*, 2020). Pour cette étude la prosodie a été aussi bien mesurée avec des paramètres acoustiques que de façon subjective. Le premier intérêt ici était d'étudier la dimensionnalité de la prosodie en prenant en compte ses dimensions subjective

et acoustique et étudier son évolution. Le modèle le plus adapté aux données est appelé modèle "corrélé à deux facteurs". Ce modèle est composé d'un premier facteur "scores et pauses" correspondant aux quatre paramètres évalués subjectivement (expressivité, phrasé, fluidité et vitesse) et à deux paramètres de rythme (fréquence et durée des pauses) et d'un deuxième facteur "intonation" correspondant aux variations de F0 et au contour intonatif. Ce modèle à deux facteurs est celui décrivant le mieux les données pour les six temps de mesure. Les auteurs concluent donc sur la multidimensionnalité de la prosodie en lecture : rythme et score subjectif d'une part et intonation d'autre part. On peut rapprocher ces résultats de ceux de Cowie *et al.* (2002) et Benjamin *et al.* (2013). Dans ces deux études, cherchant à lier évaluations subjectives et paramètres acoustiques, deux facteurs principaux ressortaient : les marqueurs temporels, ou "pauses naturelles", et l'expressivité, ou "intonation expressive". Dans ces trois études, les deux facteurs, bien que distincts, ne sont pas indépendants. Cette dernière étude renforce la théorie d'une multimodalité complexe de la prosodie en lecture. De plus, elle apporte un regard intéressant via l'étude longitudinale. En effet, cette bimodalité est présente dés le CP et semble stable dans le temps, du moins pour les trois premières années d'apprentissage. Young-Suk *et al.* (2020) ont également cherché un lien entre ces deux facteurs, le décodage et la compréhension orale. Il apparait que le facteur "scores et pauses" est principalement lié au décodage, lien se renforçant au cours du temps. On retrouve donc bien ici le rôle important de l'acquisition du décodage dans le développement de la prosodie. Ce facteur est également lié à la compréhension orale dans une moindre mesure. Le facteur "intonation" est faiblement lié au décodage et ce lien reste stable au cours du temps. Les auteurs concluent à un effet de goulot d'étranglement ("bottleneck") du décodage dans l'apprentissage de la lecture, dont la maitrise conditionne le progrès de l'expressivité. Les auteurs insistent cependant sur la nécessité de poursuivre l'étude longitudinale, cet effet pouvant être trés important dans les années scrutées ici, et se réduire avec l'acquisition de la lecture. Cette étude vient donc également renforcer l'hypothèse émise dans notre article : les liens entre prosodie, vitesse et compréhension varient au cours du développement. Durant les premières années d'apprentissage de la lecture, l'acquisition du décodage et de l'automaticité priment et influencent fortement les acquisitions des autres compétences en lecture. Une fois leur maitrise suffisante, les compétences en expressivité et compréhension peuvent s'épanouir.

Ces deux études ont donc confirmé les précédentes sur les acquisitions du CP au CE2 : une maitrise rapide des pauses et plus tardive des variations de F0, un lien fort avec le décodage pendant ces premières années. Les deux études concluent sur la nécessité d'aller plus loin, en s'intéressant aux lecteurs plus âgés.

1.2.2 La prosodie en lecture au collège et au lycée

A la date de soumission de notre article, peu d'études s'intéressaient au développement de la fluence, plus particulièrement de la prosodie en lecture, au delà du primaire. Depuis 2018, plusieurs études ont été publiées sur les lecteurs adolescents. Ces études sont d'autant plus intéressantes qu'elles prennent en compte la variabilité plus large des textes rencontrés par les collègiens et les lycéens, des textes narratifs plus complexes et surtout des textes informatifs qui demande une prise d'informations plus précise.

Une compétence toujours en développement

Les premières études s'intéressant à des élèves du second degré se sont axées autour du lien fluence-compréhension, par exemple Paige *et al.* (2014) et Yildirim *et al.* (2018). Ces études ne por-

taient donc pas spécifiquement sur les compétences en fluence d'élèves du secondaire mais plutôt sur la compréhension. De plus elles utilisent toutes l'échelle multidimensionnelle de fluence "Multidimensionnal Fluency Scale" (MDFS) (Rasinski, 2004), donc une évaluation subjective de la prosodie. Elles montrent que la prosodie continue de se développer au collège. Les études acoustiques chez les plus jeunes, présentées dans notre article, montrent que la prosodie est encore immature à la fin du cycle primaire, particulièrement au niveau des paramètres d'intensité et de mélodie, principaux véhicules de l'expressivité. Une étude récente portant sur des collégiens espagnols (Alvarez-Canizo *et al.*, 2020), utilisant des marqueurs acoustiques de la prosodie, permet de compléter ces données sur les adolescents. Cette étude porte sur les deux premières années de cycle secondaire, c'est à dire sur des élèves de 11 à 13 ans. Alvarez-Canizo *et al.* (2020) montre que le nombre d'erreurs continue de diminuer avec le niveau. Ils observent également une diminution de la durée de lecture, du pourcentage de temps alloué aux pauses et du nombre de pauses, aussi bien grammaticales que non grammaticales. Par ailleurs, on retrouve une augmentation des variations de F0 avec l'âge. La difficulté majeure rencontrée par les plus jeunes semble être le manque d'anticipation de la fin des phrases, qui impacte l'intonation. La pente de F0 en fin de phrases déclaratives est plus forte chez les plus jeunes, comme s'ils signalaient soudainement la fin de la phrase par manque d'anticipation du point final. Ces tendances sont dans la droite ligne des observations faites en primaire et confirment que face à des textes plus complexes adapatés à leur niveau scolaire, la prosodie est toujours en développement au début du cursus secondaire. Ces données viennent confirmer l'hypothèse émise à partir des évaluations subjectives existantes : à la fin du primaire la prosodie en lecture continue de se développer, y compris sur des facteurs basiques tel que le nombre de pauses et l'intonation en fin de phrase.

La confrontation à des types de textes plus nombreux et plus complexes

Une particularité de l'enseignement du secondaire dans l'apprentissage de la lecture est la transition du *apprendre à lire* vers le *lire pour apprendre* à une plus grande échelle. La lecture est alors un outil indispensable pour les apprentissages de tout type. Les textes sont plus complexes et les adolescents sont de plus en plus confrontés à des textes informatifs et de moins en moins narratifs. Or ces textes documentaires diffèrent des textes narratifs dans leur difficulté perçue (Kuhn et Schwanenflugel, 2019). Cette difficulté est majorée dans le cas de textes abordant des sujets inconnus. En effet d'aprés le modèle *Simple View of Reading* proposé par Hoover et Gough (1990), la compréhension des textes dépend des capacités de décodage, qui ne doivent plus poser de problèmes à cet âge là, du vocabulaire et surtout des connaissances préalables dans le sujet abordé. Kuhn et Schwanenflugel (2019) ajoutent à cela l'importance de la prosodie qui, selon elles, permet à l'élève de mieux comprendre ce qu'il lit, et à l'enseignant d'évaluer la compréhension de l'élève. En effet, le découpage du texte par le phrasé aide à le comprendre, entre autre via une facilitation de l'attention et de la mémoire de travail. Finalement, ces auteurs insistent sur l'importance d'évaluer une fluence situationnelle et pas uniquement à partir de tests standardisés. Un adolescent peut très bien avoir un très bon score sur un texte narratif classique mais se retrouver en difficulté dans un texte documentaire sur un sujet qu'il maitrise peu. La fluence évaluée dans ces différentes situations sera alors révélatrice des difficultés rencontrées.

Alvarez-Canizo *et al.* (2020) se sont intéressés à cette différence entre textes narratifs et explicatifs. Chez les adultes Schwanenflugel *et al.* (2016) avaient observés l'usage de pauses plus longues dans les textes explicatifs et des variations de F0 dans les textes narratifs chez les adultes. Dans cette étude chez les adolescents, Alvarez-Canizo *et al.* (2020) observent plus d'erreurs lexicales

dans les textes narratifs mais plus de pauses et de variations de F0 dans les textes explicatifs. Cette différence avec les adultes pourrait s'expliquer par une difficulté plus importante dans la compréhension des textes explicatifs ou documentaires. Les enfants de primaire ont également tendance à accentuer leur prosodie en lisant des textes complexes (Benjamin et Schwanenflugel, 2010). Alvarez-Canizo *et al.* (2020) ont également pu observer que les plus jeunes s'appuyaient plus sur la prosodie pour comprendre le texte : la compréhension dans les textes explicatifs est liée à la grammaticalité des pauses et à l'allongement des voyelles finales, chez les plus jeunes, et au temps de lecture et à la conscience syntaxique, chez les plus agés. Pour les textes narratifs, seul un lien au score de tests sémantiques apparaît chez les plus jeunes. Cette étude vient donc bien confirmer les propos de Kuhn et Schwanenflugel (2019). Chez les lecteurs du secondaire, les textes narratifs, plus familiers, sont plus facilement compris que les textes explicatifs qui sont plus complexes et abordent des sujets nouveaux. Un élève ayant un bon score de fluence sur des textes narratifs peut être en difficulté de lecture sur d'autres textes. Tous ces auteurs pointent donc la nécessité de la prise en compte de cette diversité des textes dans l'évaluation et l'enseignement de la lecture dans le secondaire, car les compétences de lecture sont toujours en développement.

1.2.3 Lier prosodie compréhension

Comme souligné dans notre article, au delà du développement de la prosodie, le lien prosodie-compréhension est un intérêt de recherche majeur. L'existence de ce lien a été montré à de nombreuses reprises, mais de nombreuses critiques persistent : la relation de causalité, la robustesse des méthodes de mesure de la compréhension et de la prosodie, l'impact de l'automaticité et des autres compétences cognitives dans les premières années... Les études continuent donc de se multiplier sur le sujet avec des approches alternatives aussi bien dans les méthodes d'évaluation que dans les sujets étudiés.

Évaluation

Traditionnellement la compréhension et la fluence sont évaluées à l'aide de test standardisés. Les tests de compréhension utilisés sont essentiellement des tests en lecture silencieuse (texte et questions). Ces tests s'adressent donc plutôt à des lecteurs accomplis et questionnent la compréhension globale hors-ligne du texte. De plus, la compréhension et la fluence sont généralement évaluées sur des textes différents.

Kocaarslan (2019) explore le lien fluence compréhension chez des CE1 turcs. Il évalue toutes les dimensions : vitesse, précision, prosodie et compréhension sur un seul et même texte adapté au niveau scolaire des enfants. La prosodie est évaluée par la MDFS et la compréhension par le rappel du texte lu. L'auteur ne fixe pas un objectif de vitesse à l'élève pour guider sa performance : dans cette étude, les enfants sont informés qu'ils devront raconter l'histoire à l'expérimentateur une fois le texte lu à voix haute. Cette consigne oriente la lecture vers la compréhension et non vers la vitesse. On retrouve donc ici chez le tout jeune lecteur l'importance de la fluence situationnelle, prônée par Kuhn et Schwanenflugel (2019) chez les lecteurs plus âgés. Dans cette étude le score de rappel est uniquement lié au score subjectif de prosodie et pas au NMCLM, pourtant mesuré sur le même texte. Ces résultats confirment l'importance de la prosodie pour la compréhension, y compris chez les plus jeunes. Ils sont particulièrement intéressants car, chez les jeunes lecteurs, la littérature va plutôt dans le sens d'une prédominance du décodage et de l'automaticité, ce qui

n'est pas le cas ici quand la compréhension est le but explicite de l'exercice. Il faut cependant noté que la MDFS mesure également la prosodie et la vitesse, on ne peut donc pas conclure à un effet de la prosodie seule. Mais cette étude vient affaiblir l'utilisation du NMCLM comme mesure large des compétences en lecture.

Il est difficile de séparer l'impact de la prosodie en lecture sur la compréhension de celui du décodage et de l'automaticité tant ces deux paramètres sont étroitement liés. Breen *et al.* (2016) ont essayé de mesurer l'impact de la prosodie sur la compréhension, indépendamment des compétences de décodage chez des lycéens américains. Pour cela, au lieu de lire un texte à voix haute les lycéens devaient lire des phrases à voix haute, en imitant un lecteur expert qu'ils entendaient. Ce protocole permet de s'affranchir du décodage en se focalisant sur la prosodie, même si celle-ci n'est pas produite, mais uniquement imitée. Les test de compréhension écrite définissent deux groupes : normo-compreneurs et faibles compreneurs. Les variations mélodiques globales sont identiques pour les deux groupes. Cependant les faibles compreneurs imitent moins bien le phrasé que les normo-compreneurs : durée des pauses, variations intonatives marquant les frontières syntaxiques mineures. Les normo-compreneurs marquent plus les frontières syntaxiques aussi bien par l'intonation que par les durées de pauses. Pour les auteurs, cela pourrait s'expliquer par une meilleure utilisation de la prosodie en lecture silencieuse par les normo-lecteurs : ceux ci marquant plus le phrasé en lecture silencieuse, ils comprennent mieux. Cette étude montre l'importance du phrasé dans la compréhension, et soulève également la question de sa perception. L'imitation plus faible du phrasé s'accompagne-t-elle du même effet en lecture orale individuelle ? Provient-elle d'un déficit de perception ou uniquement de production ?

Intérêt pour les publics particuliers

Au-delà de l'étude de lecteur "tout-venant", l'étude de public particulier, par exemple avec des pathologies, nous permet également d'en savoir plus sur ce public et sur le fonctionnement normal. Les études présentées ici s'intéressent aux enfants présentant des déficits de compréhension.

Lalain *et al.* (2014) et Suárez-Coalla *et al.* (2016) se sont intéressés aux enfants et adultes dyslexiques respectivement en français et espagnol. Ces études ont montré une faiblesse de la prosodie, notamment du rythme et de l'anticipation de la fin des phrases, chez les lecteurs dyslexiques, ce qui impacte la qualité de leur lecture, y compris chez des lecteurs dyslexiques adultes expérimentés. On retrouve des résultats similaires dans l'étude de Jordán *et al.* (2019) avec des enfants présentant des troubles du langage. On peut émettre l'hypothèse que, comme les enfants dyslexiques et ayant un trouble du langage, les faibles compreneurs présentent une prosodie appauvrie par rapport au normo-lecteurs. Groen *et al.* (2019) s'intéressent aux faibles compreneurs. Cette étude compare la perception et la production prosodique d'enfants faibles compreneurs de CM2 avec des enfants de même âge normo-compreneurs et des enfants de même compréhension mais plus jeunes. Il apparaît que les faibles compreneurs présentent un déficit aussi bien en perception qu'en production par rapport aux enfants du même âge, mais également en perception par rapport aux plus jeunes ayant les mêmes scores de compréhension. De plus, les régressions réalisées sur les différents paramètres montrent que le score de prosodie en lecture (évalué avec la MDFS) explique bien une part propre de variance des scores de compréhension, au delà du raisonnement non verbal, du décodage et du vocabulaire. Les auteurs en concluent que si le décodage et l'automaticité ont un fort impact sur la compréhension en début d'apprentissage, celui-ci diminue petit à petit, tandis que celui de la prosodie augmente. En effet, pour eux, "la prosodie permet de faciliter l'unification entre les niveaux

phonologique, syntaxique et sémantique" et donc la compréhension. La prosodie serait donc un bon indicateur du niveau de compréhension des enfants. Un défaut de perception et de production de prosodie, aussi bien en lecture orale que silencieuse, altérerait la compréhension.

Le problème de l'influence de l'automaticité et de la prosodie sur la compréhension peut également être considéré dans l'autre sens. Les lecteurs rapides comprennent-ils tous ce qu'ils lisent ? O'Connor (2018) montre qu'au-delà d'un certain seuil, le NMCLM ne prédit plus la compréhension. Par exemple à plus de 150 MCLM, l'augmentation de la vitesse n'apporte plus rien à la compréhension chez les normo-lecteurs de CM1, et pourrait même avoir un effet néfaste sur celle-ci. D'autres paramètres influent sur cette limite : l'auteur montre que chez des enfants présentant des troubles de la lecture, cette limite est beaucoup plus basse. Ils n'ont pas besoin d'atteindre une vitesse équivalente à la parole pour comprendre. Chez les CM1, une vitesse de 40 à 90 MCLM est suffisante pour atteindre une compréhension standard. L'auteur argumente donc en faveur d'un but de "vitesse suffisante" et du travail de la prosodie pour améliorer la compréhension. Ces résultats convergent avec ceux de Applegate *et al.* (2009) qui montraient que 30% des enfants du CE1 à la seconde, jugés très bons lecteurs par leurs enseignants et présentant des scores de fluence dans le quartile supérieur, avaient des scores de compréhension faibles.

En résumé, la vitesse de lecture (évaluée par le NMCLM) conditionne la compréhension dans les premières années d'apprentissage, mais son influence sur l'appréhension du sens diminue rapidement au profit d'autres compétences. La prosodie en lecture est régulièrement mise en avant dans la littérature comme "successeur" du NMCLM pour les plus grands, pour des publics présentant des troubles de la lecture et pour les plus jeunes comme un potentiel prédicteur de difficultés futures.

1.2.4 Le casse-tête de l'évaluation de la fluence incluant la prosodie

La prosodie est de plus en plus mise en avant comme un outil intéressant pour évaluer les compétences en lecture des élèves. Cependant, comme le rappelle Dowd et Bartlett (2019), la question de la prise en compte de la prosodie se pose depuis longtemps (Dowhower, 1991) sans pour autant être concrètement mise en place dans l'évaluation et l'enseignement. Selon eux, cela s'explique en grande partie par l'image de simplicité du NMCLM : simple d'utilisation, quantitatif et objectif, comparativement à une évaluation qualitative de la prosodie paraissant plus complexe. Cependant, comme on a pu le voir dans les paragraphes précédents, la situation n'est pas aussi simple : ce NMCLM est dépendant du texte et des connaissances des élèves et il n'est pas forcément révélateur des compétences en lecture des élèves. Dowd et Bartlett (2019) montrent que même la mesure de la précision par ce type d'évaluation peut être défectueuse. Tester à travers 11 langues en CE1, la limite de temps imposée dans la lecture pénalise une grande partie des enfants aussi bien en terme de compréhension que de précision. La levée de la limite de temps amène une diminution de la vitesse de lecture mais également un augmentation significative des scores de compréhension et de précision. Les enfants en difficulté et bilingues semblent par ailleurs être les plus pénalisés. Le recours à un autre type de test ayant une consigne plus qualitative laissent le temps aux enfants de s'appliquer. Rasinski *et al.* (2019) prônent également une approche qualitative de fluence. La façon d'évaluer influençant inévitablement la façon de l'enseigner, axer l'évaluation sur le NMCLM engendre une course à la vitesse, alors qu'une approche qualitative de la fluence remettrait la qualité de la lecture orale au centre de l'apprentissage, comme un moyen de communication et de compréhension des textes. Ainsi, bien que la prosodie paraisse plus complexe à évaluer, car trés subjective, son aspect qualitatif est particulièrement intéressant.

L'évaluation qualitative de la fluence incluant la prosodie passe par des échelles subjectives, de type MDFS (cf tableau 1 de l'article 1.1) parfois critiquées pour leur subjectivité ou leur utilisation obscure (Haskins et Aleccia, 2014). Initialement développée pour l'évaluation en classe, l'échelle MDFS (Rasinski, 2004) est la plus utilisée en recherche dans la littérature (Morrison et Wilcox, 2020), mais la fiabilité inter-évaluateurs est souvent avancée sans préciser les conditions d'utilisation de cette échelle et le protocole utilisé pour l'évaluation. Par exemple, certains font appel à un seul évaluateur et vérifient la fiabilité avec un deuxième évaluateur sur une vingtaine de lectures, certains évaluateurs sont des experts en lecture, d'autres des enseignants ou encore des étudiants... Moser *et al.* (2014) concluent à une bonne fiabilité de l'échelle et conseillent au moins deux évaluateurs et deux textes pour une fiabilité optimale. Smith et Paige (2019) ont réalisé une étude de généralisabilité sur la MDFS et concluent à une variance due aux évaluateurs de 0 à 2.2% et à un très bon coefficient de généralisabilité de 0.91 à 0.94. Cependant cette fiabilité et ces scores d'accords entre évaluateurs élevés cachent des protocoles différents. L'entrainement des évaluateurs, rarement mentionné dans les études utilisant l'échelle, joue un rôle clé dans ces chiffres (Haskins et Aleccia, 2014). Par exemple, cette dernière étude de généralisabilité a été réalisée avec deux jours d'entraînement et plus de 50 lectures discutées pour "modeler" la vision des évaluateurs de ce qu'est une lecture expressive. Il est peu probable que ce type de protocole soit systématiquement mis en place en recherche, comme en classe. On peut noter qu'en recherche quelques équipes utilisent des marqueurs acoustiques objectifs pour caractériser le développement des compétences prosodiques. Cependant cette méthode est complexe à mettre œuvre car elle repose sur un enregistrement des enfants et une mesure de paramètres représentatifs de la prosodie en lecture. Cela demande du matériel, une analyse automatique de la parole fiable, difficile à atteindre avec des apprentis lecteurs faisant beaucoup d'erreurs, de répétitions, d'omissions... , et le choix de paramètres représentatifs.

Ainsi, si l'utilité d'une évaluation qualitative de la lecture prenant en compte la prosodie est attestée, sa mesure reste problématique car tributaire d'évaluateurs humains ou de matériel et d'analyse complexe, et surtout de la diversité du but à atteindre aussi bien en terme d'instruction de lecture que de performance attendue. Cette question se pose de façon d'autant plus aigue dans le cadre du suivi des élèves et de leurs progrès, en étude longitudinale comme en classe. Dans ce cadre là, la question de la fiabilité dans le temps reste encore inexplorée.

1.3 Synthèse des enjeux et questionnements en suspens

Cette revue de littérature fait ressortir deux approches importantes de la prosodie en lecture. D'une part une approche cognitive questionne l'implication des compétences prosodiques dans la compréhension écrite et les étapes développementales d'acquisition de ces compétences. D'autre part, une approche plus technique vise à établir quelle méthode de mesure est la plus informante, fiable et reproductible.

D'un point de vue cognitif, on constate un intérêt important pour le lien entre prosodie en lecture et compréhension écrite. Deux dimensions de la prosodie semblent particulièrement liées à la compréhension écrite : le phrasé et l'expressivité. Les études sur les relations de causalité entre ces dimensions et la compréhension en lecture ne sont cependant pas toutes congruentes et la bidirectionnalité d'une relation évoluant avec les acquisitions est la piste privilégiée actuellement. Une meilleure connaissance du développement de la prosodie en lecture est indispensable à une meilleure compréhension la relation entre prosodie et compréhension. Au-delà d'un amélioration

des modèles de développement de la lecture et donc de la compréhension de ses mécanismes cognitifs, cette connaissance pourrait permettre de proposer un enseignement adapté, une prise en charge ciblée des enfants en difficulté de lecture et particulièrement de compréhension, et un repérage précoce des difficultés potentielles des enfants. Cet aspect est d'autant plus important qu'il n'existe pour l'instant pas ou peu d'enseignement explicite de la prosodie en lecture. Celui-ci passe essentiellement par l'écoute des modèles donnés par les parents et les enseignants.

Les quelques modèles de développement de la prosodie en lecture proposés dans la littérature nous donnent une première idée des étapes possibles : aprés l'acquisition du décodage et de l'automaticité, la mise en place du phrasé, via l'organisation des pauses, puis dans un deuxième temps celle de l'intonation linguistique, puis finalement l'intonation paralinguistique avec l'usage de variations mélodiques et d'intensité pour rendre le texte expressif. Cependant, les étapes décrites ici ne sont pas complètement dissociées les unes des autres, elles sont au contraires conjointes et se chevauchent dans le temps. On observe un manque d'études longitudinales centrées sur ce développement de la prosodie. Ces études permettraient de mieux comprendre les relations et les prérequis nécessaires au développement des dimensions de la prosodie. Finalement les résultats conduisant aux modèles présentés ici sont essentiellement issus de lecture en anglais et espagnol. On peut donc se poser la question de la transposition à d'autres langues et d'autres curriculums. De la même façon, ces modèles s'intéressent avant tout aux premières années d'apprentissage de la lecture alors que ces mêmes études montrent que la prosodie est loin d'être acquise à l'issue de ces premières années. Même chez les collégiens, ce développement n'est pas encore complet. Il est donc nécessaire de s'intéresser au développement de la prosodie au-delà des premières années. Cet aspect est d'autant plus important qu'il semble que le lien compréhension-prosodie se renforce avec le temps et la consolidation des compétences en lecture. Nous aborderons ce sujet dans les chapitres 3 et 4. L'évaluation de lecteurs plus âgés a fait naître un intérêt pour la lecture de textes explicatifs, car ce type de texte est très présent dans le secondaire et engendre plus de difficultés de compréhension chez les élèves. La prosodie peut alors être vue comme une aide à la compréhension, mais également un outil d'évaluation d'éventuelles difficultés de compréhension. Le choix des textes de travail et d'évaluation est alors crucial pour maintenir la motivation et un objectif atteignable pour les élèves. Cet aspect touche alors également à l'évaluation de la complexité des textes que nous n'aborderons pas ici, mais qui pose aussi question.

Un dernier enjeu, de taille, est la question de la mesure de la prosodie. Si la prosodie est en effet un bon indicateur des performances générales et de la compréhension en lecture, il est nécessaire de l'évaluer d'une façon fiable. Une évaluation rapide et aisée permettrait également plus facilement sa prise en compte en classe, afin d'identifier le niveau et les progrès de manière individuelle et de batir des parcours d'apprentissage différenciés. Parmi les outils actuellement à disposition, l'évaluation objective de paramètres acoustiques donne de nombreux résultats et permet de suivre précisément le développement de chaque dimension de la prosodie. Cependant elle reste compliquée à mettre œuvre car elle repose sur un enregistrement des enfants et une mesure de paramètres représentatifs de la prosodie en lecture relativement complexe et chronophage. Son utilisation reste donc limitée au laboratoire et à des cohortes relativement peu nombreuses. L'alternative est l'utilisation d'outils subjectifs où des assesseurs évaluent les performances en lecture orale. Cette technique est plus rapide et facile à mettre en œuvre. Cependant elle pose toujours question en terme de fiabilité et reproductibilité, à cause du côté subjectif de la prosodie et de l'absence d'idéal unique. Dans les deux cas, l'évaluation de la prosodie se heurte essentiellement à un problème majeur qui est l'absence d'un modèle unique. En effet, il existe de nombreuse façons de lire un texte avec expressivité, de nombreuses variations mélodiques et rythmiques différentes mais correctes. Un des principaux

enjeux de l'évaluation de la prosodie est donc de prendre en compte cette multitude de modèles licites et de l'intégrer à un outil d'évaluation fiable et simple d'utilisation. Nous aborderons ce sujet dans le chapitre 5.

Chapitre 2

Introduction à la partie expérimentale

Le travail présenté ici s'inscrit dans les problématiques présentées dans le chapitre précédent. Nous allons nous intéresser au versant cognitif de l'étude la prosodie en lecture : développement et lien avec la compréhension, tout en prenant en compte la problématique technique de l'évaluation de la prosodie.

Nous nous intéressons au développement de la prosodie, des premières années d'apprentissage jusqu'au collège, en nous posant les questions de son évaluation et son lien à la compréhension. Ce travail vise tout d'abord à reproduire en français les analyses effectuées dans d'autres langues et vérifier la convergence des résultats, en utilisant une évaluation subjective et une évaluation objective acoustique. Mais nous souhaitons également aller plus loin en nous intéressant à des sujets encore pas ou peu abordés dans la littérature, tels que la coordination lecture-respiration ou la proposition d'un outil d'évalution prenant en compte les multiples possibilités de production correctes et expressives. Ainsi notre objectif est d'aborder le développement de la prosodie et son lien avec la compréhension écrite en utilisant trois types de mesure : une évaluation subjective, une mesure objective des marqueurs acoustiques des pauses et une mesure objective prenant en compte la multiplicité des productions acceptables.

Le travail présenté ici est organisé en trois chapitre correspondant aux trois méthodes de mesures de la prosodie choisie. Dans chacun de ces chapitres nous adopterons la même approche cognitive en utilisant une mesure différente. Cette partie présente les objectifs de ces trois chapitres et les études menées pour y répondre puis les choix méthodologiques réalisés.

2.1 Présentation des études

2.1.1 Caractérisation du développement de la prosodie par évaluation subjective

La première partie de ce travail de thèse consiste tout d'abord à valider une échelle subjective de mesure de la prosodie en lecture pour le français, adaptée de l'échelle anglaise MDFS. L'évaluation du corpus transversal avec cette échelle nous permet de caractériser le développement de la prosodie, ainsi que le lien prosodie-compréhension du CE1 à la 5e.

– **Étude 1 : Étude transversale du développement de la fluence en lecture incluant la prosodie**
Dans cette étude, nous adaptons la MDFS (Rasinski, 2004) en français pour évaluer les lectures de 295 enfants du CE1 à la 5e. Nous montrons la validité de cette échelle avec plusieurs textes. Nous complétons avec la comparaison des scores donnés par différents types d'évaluateurs : experts, enseignants et étudiants. Finalement nous regardons l'évolution des scores en fonction du niveau scolaire. Nous nous attendons à une augmentation du score global et de toutes les dimensions quand le niveau scolaire augmente et à un plafonnement des dimensions de vitesse et de décodage après les premières années d'apprentissage. Nous présentons un étalonnage de cette échelle pour les textes utilisés.

– **Étude 2 : Étude du lien entre évaluation subjective de la prosodie et la compréhension écrite**
Dans cette étude nous utilisons les scores subjectifs de l'étude 1 pour établir un lien avec la compréhension écrite. Nous recherchons les prédicteurs des scores prosodiques et de compréhension et l'effet du niveau scolaire sur ces prédicteurs. Nous prenons en compte le score de fluence et des scores spécifiques aux dimensions prosodiques, ainsi que différents types de compréhension présentés dans le paragraphe 2.2. Nous nous attendons à trouver un lien prosodie-compréhension, se renforçant quand le niveau scolaire augmente, conformément à la littérature.

2.1.2 Caractérisation du développement de la prosodie par l'analyse des pauses et de la respiration

La deuxième partie de ce travail est de caractériser le développement de la lecture en étudiant des données acoustiques. Nous nous intéressons ici particulièrement aux pauses. Il s'agit d'une part de retrouver le schéma d'acquisition des pauses observées en anglais et en espagnol, présenté dans le premier chapitre, mais également les liens entre nombre, durée et grammaticalité des pauses avec d'une part notre perception de la fluence et d'autre part avec la compréhension. Nous souhaitons également amener de nouvelles connaissances sur les pauses. En effet, dans la littérature les pauses sont traitées comme un paramètre unique. Cependant elles assurent de multiples fonctions. En premier lieu, elles permettent de reprendre sa respiration pendant la lecture, de marquer la syntaxe du texte pour faciliter sa compréhension et de décoder le mot suivant. Le placement correct des pauses nécessite leur anticipation. Nous nous intéressons donc aux détails des pauses pendant la lecture, pour explorer le développement de leur planification, qu'elles soient respiratoires ou syntaxiques, pour comprendre quel impact elles ont sur notre perception subjective des lectures et pour évaluer si le lien entre pauses et compréhension dépend du type de pauses.

– **Étude 3 : Pauses et respiration du CE1 à l'adulte expert**
Cette étude vise à caractériser le développement des schémas de pauses et de respiration pendant la lecture et donc de la planification de celles-ci. Pour cela, nous analysons les pauses produites dans 295 lectures d'enfants du CE1 à la 5e. Nous nous intéressons à la distribution de la durée des pauses. Nous nous attendons à une bimodalité : pauses courtes et longues. Nous étudions l'évolution de cette distribution et les nombres, durées et grammaticalité des pauses courtes et longues. L'étude des pauses non grammaticales et de leur liens avec la ponctuation, nous renseigne sur l'évolution de la planification des pauses syntaxiques. Le même travail est réalisé pour les pauses respiratoires. Nous nous attendons à une diminution

du nombre de ces pauses avec le développement physiologique et à une diminution du nombre de pauses respiratoires non grammaticales et du délai inspiration-phonation avec le niveau, ce qui indiquerait une meilleure programmation des pauses respiratoires.

- **Étude 4 : Marqueurs acoustiques des pauses et perception de la fluence**
  Dans cette étude, nous cherchons à mettre en évidence les marqueurs acoustiques de pauses influençant le plus notre perception de la fluence d'une lecture. Pour cela nous utilisons le corpus transversal pour mettre en relation les scores subjectifs obtenus dans l'étude 1 et les paramètres acoustiques les plus saillants de l'étude 3. Nous nous attendons à trouver un effet du nombre de pauses et de pauses non grammaticales.
- **Étude 5 : Marqueurs acoustiques des pauses et compréhension écrite**
  Dans cette étude, nous utilisons les paramètres de pauses les plus significatifs du développement de la prosodie, identifiés dans l'étude 3, et modélisons leur relation à la compréhension. Nous regardons comment cette relation évolue avec le niveau scolaire, au delà du vocabulaire, de la compréhension orale et de la vitesse de lecture. Nous nous intéressons à la fois à la compréhension globale et à la compréhension en-ligne qui participe à la vitesse de lecture. Nous nous attendons à une corrélation forte avec le nombre de pauses longues et non grammaticales dans les premières années, conformément à la littérature.

2.1.3 Caractérisation du développement longitudinal de la prosodie par évaluation automatisée

Dans cette partie expérimentale, l'objectif est d'utiliser un autre type de mesure fiable, facile d'utilisation, présentant une bonne reproductibilité et permettant une utilisation dans le cadre d'un suivi longitudinal d'élèves. L'outil présenté ici nous permet de caractériser le développement de chacune des dimensions à l'EMDF et la causalité et l'évolution des liens entre ces dimensions et avec la compréhension écrite.

- **Étude 6 : Développement longitudinal de la prosodie mesurée par évaluation automatisée de la fluence**
  Dans cette étude, nous présentons un outil automatique permettant de prédire les scores à l'EMDF pour des enfants du CE1 à la 5^{e}. Cet outil se base sur les lectures de 20 lecteurs experts définissant un cadre de référence des possibilités prosodiques, et sur plusieurs textes. Nous utilisons cet outil pour évaluer des lectures dans le cadre du suivi longitudinal d'élèves testés en CE1, CE2 et CM1 et proposer un modèle de croissance pour chacune des dimensions de l'EMDF
- **Étude 7 : Etude longitudinale du développement des dimensions prosodiques de la lecture et de la compréhension écrite**
  Dans cette dernière étude, nous utilisons les scores obtenus sur les trois ans de suivi longitudinal avec l'outil présenté dans l'étude précédente, pour mettre en évidence la causalité et les variations temporelles dans les liens entre les quatre dimensions de la fluence et la compréhension écrite. Au vu de la littérature, nous nous attendons d'une part à des liens allant des dimensions d'automaticité vers les dimensions prosodiques et d'autre part à un lien de la compréhension vers la prosodie, s'inversant avec le temps.

2.2 Précisions méthodologiques

Les sept études ci-dessus reposent sur les mêmes corpus d'enregistrement de lectures d'enfants et les mêmes tests cognitifs. Les participants, le matériel et les protocoles utilisés sont présentés dans ce paragraphe. Les expérimentations ayant permis de constituer les corpus utilisés ont été approuvées par le CERGA (Comité d'Éthique de la Recherche Grenoble Alpes) par l'avis CERNI-Avis-2018-03-06-2 et déclarées à la CNIL. Nous avons également obtenu l'aval de la Directrice Académique des Service de l'Éducation Nationale de l'Isère, des Inspectrices de l'Éducation Nationale des circonscriptions concernées, des directrices des écoles et du principal du collège, pour enregistrer les enfants dans les établissements pendant le temps scolaire. Tous les enregistrements ont été réalisés avec l'autorisation écrite des parents et l'accord verbal des enfants.

2.2.1 Participants

Les enfants participants à l'étude sont scolarisés dans deux écoles et un collège de la région grenobloise, soit 14 classes. Ces établissements ne sont pas situés en Réseau d'Éducation Prioritaire et accueillent un public de milieu socio-économique diversifié.

Corpus transversal (utilisé dans les études 1 à 6)

Les données ont été récoltées auprès d'enfants du CE1 à la 5^{e} en fin d'année scolaire, en mai et juin. Au total, les autorisations parentales ont été distribuées à 372 élèves. Vingt familles ont refusé, 28 n'ont pas rendu l'autorisation. Finalement, 324 enfants ont participé aux passations. Un enfant a souhaité arrêter avant la fin. Nous avons donc obtenu les enregistrements de 323 enfants, dont 77 enfants bilingues et 7 enfants primo-arrivants (moins d'un an en France). Pour les études présentées ici, nous nous sommes focalisés sur des enfants normo-lecteurs, dont au moins une langue maternelle était le français. Nous avons donc écarté de l'analyse les données des enfants avec un diagnostic de dyslexie déclaré (N = 4) et les enfants bilingues ne parlant pas français à la maison (N = 24), soit 28 enfants. Les résultats moyens des 295 enfants restants aux tests de lecture NMCLM (Maeder *et al.*, 2018) et de raisonnement non-verbal (Raven *et al.*, 2003) sont conformes aux normes d'âge fournies avec ces tests. Les moyennes et écart-types des scores obtenus à chaque niveau scolaire sont présentés dans le tableau 2.1.

Vingt adultes ont également été enregistrés sur les mêmes textes. Ces adultes sont des enseignants et des doctorants, enregistrés au laboratoire. Ils ont tous été jugés comme lecteurs experts lors des évaluations subjectives réalisées dans le chapitre 4. Les caractéristiques de la cohorte des 295 enfants et 20 adultes sont présentées dans le tableau 2.1.

Corpus longitudinal (utilisé dans les études 6 et 7)

Les données de ce corpus correspondent aux données des CE1 du corpus transversal enregistrés au mois de mai 2018. Ces élèves ont ensuite été de nouveau enregistrés deux fois, à huit mois d'intervalle, c'est à dire en février de l'année de CE2 et octobre de l'année de CM1. Une soixantaine d'élève a été enregistrée chaque année. En prenant en compte les arrivées, départs et refus d'autorisation des parents, 49 élèves ont été enregistrés trois fois, 50 sur les années une et deux et 53 sur les années deux et trois. Comme pour le corpus transversal, le corpus longitudinal ne prend en compte que les enfants sans diagnostic de dyslexie et de langue maternelle française. Les caractéristiques de la cohorte sont présentées dans le tableau 2.2.

TABLE 2.1 – Description des participants au corpus transversal par niveau

	CE1	CE2	CM1	CM2	6e	5e	adulte
nombre de sujets	61	47	54	63	33	37	20
filles	27	25	26	23	15	21	10
nombre de classe	4	3	4	4	2	2	NA
âge moyen	7 : 11(3.8)	8 : 11(3.4)	9 : 11(5.1)	10 : 11(4.5)	11 : 11(5.4)	12 : 8(5.4)	29 : 6(85)
étendue	7 : 2 − 8 : 5	8 : 5 − 9 : 5	8 : 5 − 10 : 5	10 : 3 − 12	11 : 4 − 13	11 : 8 − 13 : 9	
NMCLM	116(43)	160(62)	199(59)	237(84)	232(59)	287(76)	
PM38 (scores bruts)	33.6(9.3)	37.5(8)	39.7(6.8)	43.7(6.8)	38.6(5.9)	41.5(6.3)	NA
PM38 (scores refs)	33	36	40	41	38.9(6.4)	42(7.1)	NA

Note : NMCLM = mots correctement lus en 2 minutes aux tests d'Evaléo , PM38 : scores bruts et de références aux matrices de RAVEN, 50e percentile pour les primaires, moyenne et écart-type pour les collégiens

TABLE 2.2 – Description des participants au corpus longitudinal

année	1	2	3
niveau	CE1	CE2	CM1
nombre de classe	4	4	6
nombre de sujets	51	61	55
nombre de filles	21	24	22
âge moyen	7 : 11(3.8)	8 : 7(3.8)	9 : 1(3.8)
étendue	$7:2-8:5$	$7:10-9$	$8:9-9:9$
NMCLM	116(43)	175(41)	219(49)

Note : Age en années et mois, écart-type et étendue en mois. NMCLM : nombre de mots lus en 2 min aux test Evaléo

TABLE 2.3 – Tests écrits réalisés pour chaque corpus pour toutes les années d'enregistrement

tests	corpus transversal	corpus longitudinal	
		année 2	année 3
prosodie	X	X	X
NMCLM	X	X	X
CE globale	X	X	X
CE locale		X	X
CO	X	X	X
vocabulaire	X		
PM38	X		

Note : - CE : compréhension écrite - CO : compréhension orale - PM38 : raisonnement non-verbal

2.2.2 Matériel

Le tableau 2.3 récapitule les tests menés conjointement aux enregistrements des corpus transversal et longitudinal.

Textes pour l'évaluation de la prosodie

Pour ces deux corpus, les enfants ont lu deux textes. Ces deux textes ont été écrits spécialement pour ces études, à partir de la nouvelle "Soupçon" de B. Friot (Histoires Pressées, 2007). Le texte 1 est un texte narratif de 174 mots. Ce texte a été écrit pour permettre d'étudier le phrasé. Il comporte quatre phrases courtes de sept à neuf mots, comportant une seule proposition, et huit phrases longues de 15 à 17 mots, comportant deux ou trois propositions. Parmi les phrases longues, la moitié contient deux virgules, l'autre moitié ne contient aucun signe de ponctuation. Les phrases longues, avec et sans ponctuation, ont des fréquences lexicales moyennes équivalentes avec 4,528 occurrences pour un million de mots (SD = 325) pour les phrases avec ponctuation et 4,406 occurrences pour un million de mots (SD = 340) pour les autres. Ces fréquences ont été calculées à partir des données de Manulex (Lété *et al.*, 2004). Cette configuration permet de varier les possibilités de pauses et de mettre en avant l'utilisation de la ponctuation et le respect de la syntaxe dans le schéma des pauses.

Le texte 2 est un dialogue en lien avec le texte précédent pouvant être lu indifféremment avant ou après ce dernier. Il a été écrit pour étudier le développement de l'expressivité et plus

particulièrement les variations d'intonation. Il contient neuf phrases déclaratives, huit phrases interrogatives et quatre phrases exclamatives. Le choix du dialogue visait à induire l'expressivité. Il est en effet parfois difficile d'ajouter beaucoup d'expressivité sur les textes narratifs.

La consigne associée à ces deux textes est de les lire à voix haute "comme si tu lisais une histoire à un petit de maternelle". Cette consigne est donnée une fois encore pour encourager l'enfant à lire pour quelqu'un d'autre et non pas vite, comme certains en ont l'habitude, et ainsi induire l'expressivité. Ces deux textes sont présentés en annexe A.

Vitesse de lecture NMCLM

La vitesse de lecture est évaluée à l'aide du test de "La Mouette" tiré de la batterie d'évaluation des troubles du langage Evaléo 6-15 (Maeder *et al.*, 2018). Ce test de fluence est initialement destiné au dépistage des troubles d'apprentissage de la lecture. Ce test étant étalonné, il nous permet de vérifier que nos sujets sont bien normo-lecteurs et ne présentent pas de troubles avérés. Il est composé de deux textes appariés en syntaxe et fréquence lexicale. Le premier, intitulé "La Mouette", est un texte signifiant de 450 mots. Le deuxième est un texte non signifiant, comportant exactement la même syntaxe et ou chaque mot a été remplacé par un autre de même classe grammaticale et de même fréquence lexicale. Ce texte est donc syntaxiquement correct mais sémantiquement absurde. Lors de ce test, les enfants lisent d'abord le texte non signifiant. La consigne est de lire pendant deux minutes en essayant d'aller le plus loin possible dans le texte. Les enfants sont prévenus que ce texte n'a aucun sens et qu'il ne faut pas chercher à comprendre quelque chose, juste lire les mots qui sont écrits. Cet avertissement évite que les enfants, ne comprenant pas, bloquent et arrêtent de lire. Pendant la lecture, l'expérimentateur note les omissions et les erreurs de lecture. Quand un enfant répète un mot, il est compté comme correct si sa dernière répétition est correcte. Au bout de deux minutes, l'expérimentateur note le dernier mot lu. Les enfants lisent ensuite le texte signifiant selon le même protocole. Ils sont prévenus en amont que ce deuxième texte raconte une histoire. Ce test permet de calculer pour chacun des textes, signifiant et non signifiant, le nombre de mots lus, le nombre d'omissions, le nombre de mots correctement lus et le pourcentage de mots correctement lus. Ce test permet donc d'obtenir un score d'automaticité, via la vitesse de lecture, et de précision, via le pourcentage de mots lus, pour un texte signifiant et un texte non signifiant.

Vocabulaire

Le test de vocabulaire utilisé est un questionnaire à choix multiples provenant de l'étude ORTHOLEARN (Reilhac *et al.*, 2018). Ce test comprend 30 items de difficulté croissante. Pour chaque item, un mot de référence est proposé accompagné de cinq propositions. Les enfants doivent choisir le mot ayant le même sens que le mot référence. Le synonyme est présenté parmi les cinq propositions au côté de quatre distracteurs : phonologique, sémantique, mot ayant un lien de classe et un n'ayant aucun lien avec le mot référence. La réponse correcte rapporte un point, le mot ayant un lien de classe un demi point. Ce test est réalisé en passation collective en classe. Après la passation des consignes en groupe, deux exemples sont réalisés en commun. Chaque item (mot référence et propositions) est lu à haute voix par l'examinateur, pour éviter les erreurs de lecture. Les enfants ont ensuite cinq secondes pour cocher leur réponse. Ils avancent donc tous au même rythme et n'ont pas la possibilité de revenir en arrière. Ce test est présenté en Annexe B.

Raisonnement non-verbal

Le raisonnement non-verbal est mesuré avec le test des matrices de RAVEN (Raven *et al.*, 2003). Ce test est composé de cinq séries de douze items, chaque série est ordonnée par difficulté croissante. Les cinq séries sont également de difficulté croissante. Pour chaque item une matrice à trou est présentée avec six à huit propositions pour la compléter. Les enfants doivent reporter sur la feuille réponse le numéro de la proposition pouvant compléter la matrice. Ce test est réalisé en passation collective en classe, chaque élève est doté d'un porte-vue contenant la totalité des items. Les consignes sont données en groupe, deux exemples sont faits en commun. Ensuite les enfants réalisent seuls, à leur rythme, la totalité du test. Les élèves de primaire disposent d'un temps illimité pour réaliser l'ensemble de la tâche. En pratique la passation dure de 30 à 40 minutes suivant les niveaux. A partir de la 6^{e}, le temps est limité à 20 min afin d'être dans les conditions de l'étalonnage.

Compréhension orale

Pour le corpus transversal, le test de compréhension orale est le test "Sacré Nestor" (Crunelle *et al.*, 2000), étalonné chez les enfants de 7 à 11 ans pour diagnostiquer d'éventuels problèmes de compréhension verbale. Le test est composé d'une histoire sous forme d'enregistrement audio et de douze questions ouvertes sur cette histoire. La passation a lieu collectivement, en classe, avec réponses écrites aux questions. Les enfants sont informés qu'ils vont entendre une histoire une fois et devront ensuite répondre à des questions sur cette histoire. Ils écoutent ensuite l'enregistrement de l'histoire de Nestor. L'enregistrement dure quatre minutes. Les questions sont ensuite lues par l'examinateur et les enfants y répondent individuellement, par écrit, au fur et à mesure de leur lecture. Ceci afin d'éviter toutes erreurs de lecture dans les questions. On comptabilise ensuite le nombre de réponses correctes.
Pour le temps 2 du corpus longitudinal, le test de compréhension orale de texte de la batterie Evaleo 6-15 a été utilisé (Maeder *et al.*, 2018). Dans ce test, les enfants entendent l'enregistrement d'un paragraphe court (environ une minute) à deux reprises. Ils doivent ensuite répondre à l'écrit à six questions ouvertes puis numéroter six phrases en ordre chronologique. Les consignes et le protocoles sont les mêmes que précédemment. Le score est comptabilisé sur 12, avec un point par réponse correcte aux questions et un point par phrase correctement placée dans l'ordre chronologique.

Compréhension écrite

La compréhension écrite a été mesurée à différents niveaux. La première mesure est une mesure traditionnelle de compréhension de texte. Elle se compose d'un texte avec des questions littérales et d'inférence. Il s'agit plutôt d'une mesure de compréhension globale du texte. La deuxième mesure est une mesure plus locale. Il s'agit d'un texte à trou avec des propositions. Finalement, en utilisant les scores au test de vitesse, il est possible d'évaluer l'apport de la compréhension à la vitesse de lecture. Nous appelons ce score la compréhension en ligne.

Compréhension globale Les tests de compréhension écrite globale se composent d'un texte narratif de difficulté correspondant au niveau scolaire des sujets et de dix questions littérales et d'inférences. Trois textes ont été utilisés pour le corpus transversal : "Buddha" (CE1/CE2), "Le monde d'en haut" (CM1/CM2) et "Dick Sand" (6^e/5^e). Les CE1 ont travaillé sur un texte supplémentaire (Le Pêcheur). Ce texte plus court et sans inférence leur a été proposé en premier pour prendre en compte les lecteurs en difficulté qui pourraient abandonner face à un texte trop long. Tous ces textes sont issus de l'étude LONGIT (Bressoux *et al.*, 2016) et sont joints en annexe C. La passation collective s'est déroulée en classe. La consigne était de lire le texte et de répondre aux questions en essayant d'être précis. Les enfants pouvaient revenir autant que nécessaire au texte pour répondre aux questions. Le temps de passation était limité à 25 min, ce qui était suffisant pour que tous les enfants finissent. Le score correspond au nombre de réponses correctes.

Pour le corpus longitudinal, deux textes étaient proposés à chaque temps de mesure. En CE2, les enfants ont travaillé de nouveau sur Buddha, déjà fait en CE1, et sur "Les 7 corbeaux". En CM1, les enfants ont de nouveau travaillé sur "Les 7 corbeaux" et sur "Destin croisés". Ces deux derniers textes proviennent également de l'étude LONGIT (Bressoux *et al.*, 2016). Le protocole était le même que pour le corpus transversal. Tout le matériel est présenté en annexe **??**.

Compréhension locale La compréhension locale est mesurée en utilisant un "Maze test". Ce type de test, peu utilisé en français, est largement utilisé dans les pays anglophones pour mesurer la compréhension écrite (Hosp *et al.*, 2016). Il s'agit d'un texte labyrinthe. Tous les sept mots, un mot est remplacé par trois propositions : le mot attendu, un distracteur syntaxique et un distracteur sémantique. L'enfant doit lire le texte en entourant au fur et à mesure les mots attendus à chaque fois que plusieurs solutions sont proposées. Il n'est donc pas nécessaire d'avoir compris le sens général du texte pour réussir ce test. Une compréhension locale au niveau de la phrase ou de quelques phrases à la suite, permet de répondre correctement.

Les tests utilisés ici ont été créés pour ces études à partir de textes de littérature jeunesse. Les textes originaux ont été modifiés pour avoir une variété grammaticale dans les choix multiples. Ils sont présentés en annexe C. Le test est réalisé en passation collective en classe. Deux phrases exemples sont corrigées en grand groupe pour vérifier que tout le monde a compris la consigne, puis les enfants réalisent le test seul. Le temps de passation est limité à 20 min, ce qui était suffisant à tous les enfants pour finir le test. Le score obtenu correspond au nombre de réponses obtenues.

Compréhension en-ligne Nous avons utilisé un score de compréhension en-ligne. Ce score est calculé en faisant la différence entre le nombre de mots correctement lus entre les textes signifiant et non-signifiant utilisés pour mesurer la vitesse de lecture (Maeder *et al.*, 2018). En effet, cette différence est un bon indicateur de la façon dont les enfants s'appuient sur leur compréhension du texte pour anticiper la suite et ainsi lire plus rapidement. Il faut cependant noter que dans certains cas, cette différence peut-être négative : ces enfants lisent plus lentement quand ils comprennent. Cela peut s'expliquer par une différence de stratégie. Ce sont souvent des lecteurs moyens qui veillent à bien comprendre ce qu'ils lisent et prennent donc le temps de bien comprendre le texte signifiant et de se corriger si besoin, alors qu'ils lisent le texte non-signifiant sans se poser de question ou essayer de se corriger en cas d'erreurs.

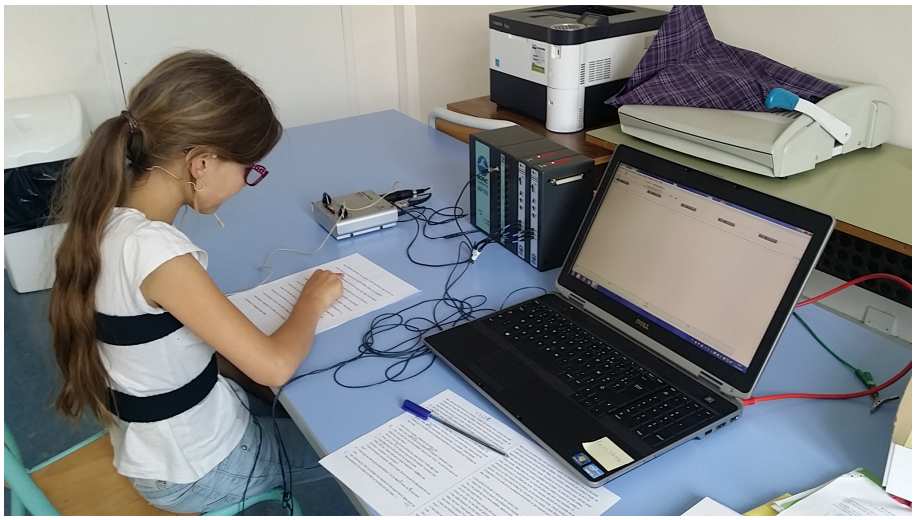


FIGURE 2.1 – Organisation matérielle des passations individuelles

2.3 Protocole de passation

Pour la récolte des données, les élèves participaient à deux passations : une passation individuelle, pour la récolte des données audio et respiratoires, une passation collective pour la récolte des données écrites. La récolte de données a été effectuée dans les écoles pendant le temps scolaire.

2.3.1 Passation individuelle

Les passations individuelles se déroulent dans une pièce calme, à proximité directe de la salle de classe. Pendant cette passation, il est demandé aux enfants de lire cinq textes. Leur voix et leur respiration sont enregistrées pendant les lectures.

L'installation d'enregistrement comprend un microphone-casque Schur Beta 53 relié à un amplificateur Behringer MIC100 pour l'enregistrement de la voix et deux ceintures RESP150 placées autour de l'abdomen et du thorax pour enregistrer les mouvements respiratoires. Les deux dispositifs sont reliés à un Biopac MP150 qui assure la synchronisation des signaux. L'acquisition est réalisée avec le logiciel Acqknowledge. Le dispositif est présenté en figure 2.1. L'utilisation d'un micro casque permet d'avoir une distance bouche-micro stable et d'éviter que les enfants parlent directement dedans, ce qui améliore la qualité du signal et évite les saturations.

Le protocole de passation suit les étapes suivantes :

- Accueil. Vérification de l'identité et du numéro d'anonymat. Présentation du matériel et explication sur le déroulement de la séance. Les enfants sont informés qu'ils peuvent arrêter quand ils le souhaitent.
- Installation et réglage. Installation des ceintures et du micro-casque. On vérifie que les ceintures ne gênent pas les enfants. Ils lisent ensuite quelques lignes d'un premier texte, pour vérification des bons réglages des ceintures et du volume. On leur présente également les signaux enregistrés sur l'ordinateur.
- Lecture des textes 1 et 2. Ces deux textes sont proposés dans un ordre aléatoire. Les enfants n'ont pas la possibilité de lire le texte en silence avant l'enregistrement. Ils découvrent donc le texte lors de l'enregistrement. La consigne est : " Lis ce texte comme si tu lisais une histoire à des petits de maternelle".

- Lecture des textes Evaleo. Les enfants lisent d'abord le texte non signifiant puis le texte signifiant. La consigne est : " Tu vas lire pendant deux minutes, il faut que tu essayes d'aller le plus loin possible dans le texte, en essayant de ne pas trop faire d'erreurs. " Ils sont prévenus à l'avance que le texte non signifiant n'a pas de sens et qu'il ne faut pas chercher à comprendre.
- Retrait des ceintures et du micro-casque. Réponses aux éventuelles questions.

2.3.2 Passation collective

Les passations collectives se déroulent en classe en présence de l'enseignant et de l'expérimentateur, pour les classes simple niveau ,et en petit groupe dans une salle à proximité de la classe en présence de deux expérimentateurs, pour les classes double-niveau. Dans les deux cas ces passations regroupent tous les tests écrits présentés ci-dessus. Elles sont organisées en séances de 40 à 50 minutes et comprennent un à trois exercices. Pour chaque test, les consignes sont données en grand groupe. A leur suite, un exemple est fait collectivement, puis un exemple est fait individuellement puis corrigé collectivement. Une fois les réponses aux éventuelles questions apportées, chaque enfant réalise l'exercice individuellement. Les enfants ont la possibilité de poser des questions, la réponse est fournie dans la mesure où elle ne fausse pas le test. Il est proposé aux enfants ayant des problème d'écriture (dysgraphie, lenteur, dysorthographie) une aide dans l'écriture des réponses sous forme de dictée à l'adulte. Les tests sont répartis comme suit :

- Corpus transversal
  - Séance 1 (50 min) : compréhension écrite globale (25 min), compréhension orale (15 min), vocabulaire (10 min)
  - Séance 2 (45 min) : RAVEN
- Corpus longitudinal
  - Séance 1 (40 min) : compréhension écrite globale (25 min) + compréhension écrite locale (15 min)
  - Séance 2 (40 min) : compréhension écrite globale 2 (25 min) + compréhension orale (15 min)

Chapitre 3

Caractérisation du développement de la prosodie par évaluation subjective

Dans ce premier chapitre expérimental, nous allons aborder le développement de la prosodie en utilisant une échelle d'évaluation subjective. Comme vu dans le chapitre 1, une grande majorité des études s'intéressant à la prosodie en lecture utilise des échelles d'évaluation subjective de la fluence. La prosodie est alors intégrée comme une dimension de la fluence au même titre que la précision et le décodage. C'est à partir de ce type d'évaluation que le lien prosodie-compréhension a été mis en évidence dans la littérature.

Il est cependant important de noter que ces échelles sont des échelles de fluence, et non de prosodie. Celle-ci est intégrée à la fluence, en général via des dimensions de phrasé et d'expressivité. Nous proposons donc dans ce chapitre d'utiliser une échelle de fluence, traduite et adaptée de l'anglais pour le français, pour évaluer la prosodie en lecture et caractériser son développement subjectif. Dans une première étude, nous utiliserons l'évaluation subjective du corpus transversal pour proposer des étapes de développement du CE1 à la 5^{e}. Dans une deuxième étude, nous utiliserons ces mêmes données pour retrouver le lien entre score à l'échelle de fluence et compréhension, mentionné dans le chapitre 1. Nous irons plus loin en étudiant également le lien entre la compréhension écrite et chacune des dimensions prosodiques de la fluence : phrasé et expressivité.

3.1 Étude 1 : Étude transversale du développement de la fluence en lecture incluant la prosodie

Afin de pouvoir étudier le développement de la prosodie en lecture français d'un point de vue qualitatif, nous avons tout d'abord traduit et adapté une échelle existante, la MDFS (Rasinski, 2004) en français. Cette échelle a ensuite été utilisée pour évaluer les lectures des deux textes du corpus transversal. Différents groupes d'évaluateurs sont intervenus : enseignants, chercheurs et étudiants. Les évaluations ont permis d'une part de montrer la validité de cette échelle et d'en proposer un étalonnage, et d'autre part de caractériser le développement de la prosodie en lecture d'un point de vue qualitatif.

3.1.1 Présentation de l'échelle

La traduction et l'adaptation de la MDFS a fait l'objet d'un article accepté pour publication en février 2020 dans la revue l'Année Psychologique.

Godde, E., Bosse, M. L., & Bailly, G.. Echelle Multidimensionnelle de Fluence : nouvel outil d'évaluation de la fluence en lecture tenant compte de la prosodie, étalonnée du CE1 à la 5e. *L'Année Psychologique.*(sous presse)

Le lien entre prosodie en lecture et compréhension écrite, régulièrement mis en évidence à différents niveaux scolaires et dans différentes langues, attire de plus en plus d'attention sur l'évaluation de la fluence. En effet, celle-ci est traditionnellement évaluée par le nombre de mots correctement lus par minute, ce qui prend en compte précision et automaticité, mais omet la prosodie, expressivité et phrasé. Cette mesure tend à entretenir la confusion entre fluence et vitesse, ce qui peut être nuisible à la compréhension. Les dimensions prosodiques de la fluence peuvent être évaluées à l'aide d'indicateurs acoustiques : pauses, fréquences de variations et taux de prononciation pour les marqueurs rythmiques, amplitude et durée des variations de mélodie et d'intensité pour l'expressivité. Ce type de mesure est surtout employé pour des études développementales. Les dimensions prosodiques de la fluence peuvent également être évaluées au moyen d'échelles subjectives. Cet article présente une échelle de ce type pour le français, traduite et adaptée d'un outil existant en anglais. 295 enfants du CE1 à la 5e, normo-lecteurs et de langue maternelle française, ont été enregistrés pendant leur lecture de deux textes. Ces lectures ont été évaluées par trois juges en utilisant une adaptation de la MDFS (Rasinski, 2004) en français, que nous avons intitulé Echelle MuldiDimensionnelle de Fluence (EMDF). Cette échelle comprend quatre dimensions : expressivité, phrasé, décodage et vitesse. Après un entraînement sur une quinzaine de lectures, les juges écoutent la première minute de lecture et donnent un score entre 1, non maitrisé, et 4, maitrise experte, pour chaque dimension. Dans l'étude présentée ici, les coefficients d'accord inter-évaluateurs sont bons, avec des variations, sur chacune des dimensions, et trés bons pour le score global, ce pour les deux textes. Cet article présente des scores globaux de référence par niveau pour chacun des textes. La sensibilité de l'EMDF est bonne jusqu'en CM2, au-delà les scores plafonnent, mais l'EMDF garde cependant une bonne sensibilité pour les élèves de collège en difficulté. Les scores à l'EMDF sont fortement corrélés au NMCLM mesurée avec "La Mouette" (Maeder *et al.*, 2018) pour les CE1-CE2. Cette corrélation diminue ensuite significativement avec le niveau scolaire. Cette diminution montre bien que cet outil apporte des informations supplémentaires dans l'évaluation de la lecture, par rapport à ceux utilisés actuellement. Les scores à l'EMDF sont également corrélés à des mesures automatiques de paramètres verbaux et prosodiques, ce qui renforce sa validité. Finalement, à travers cette étude, l'EMDF s'est révélée être un outil sensible, fidèle et valide pour évaluer la fluence en français, en prenant en compte ses dimensions prosodiques. Elle permet une évaluation qualitative, large, apportant des informations pertinentes pour évaluer plus globalement les compétences en lecture, par exemple en permettant de révéler des stratégies différentes chez les jeunes lecteurs, en s'intéressant aux différentes dimensions. Son utilisation requiert un protocole particulier à respecter, notamment dans le cas d'évaluateurs multiples, pour préserver une bonne fidélité de l'outil. Son utilisation avec les textes présentés dans cet article en fait un bon outil d'évaluation du CE1 au CM1. Au-delà il peut permettre le repérage de lecteurs en difficulté. La grille d'évaluation peut cependant être utilisée avec d'autres textes adaptés au niveau scolaire des lecteurs.

Échelle Multi-Dimensionnelle de Fluence : nouvel outil d'évaluation de la fluence en lecture prenant en compte la prosodie, étalonné du CE1 à la 5e

Résumé

Dans le contexte scolaire, la fluence de lecture des enfants est généralement évaluée par un texte à lire "le plus vite possible", ce qui tend à entretenir une confusion entre vitesse et fluence. Or, lire vite n'est pas suffisant pour être un bon lecteur. Au-delà de la vitesse, du décodage et de l'automatisation, la prosodie, c'est à dire le phrasé et l'expressivité, est également une compétence indispensable à une lecture fluente et à la compréhension du texte lu. Nous proposons ici un outil d'évaluation de la fluence en lecture incluant des dimensions de jugement de la qualité de la prosodie, de la contribution au phrasé et à l'expressivité des productions orales. Cet outil a été construit à partir d'un outil existant en anglais et adapté à la langue française. Les résultats de lecture de 295 enfants du CE1 à la 5e sont présentés en guise d'étalonnage. L'analyse des cotations effectuées avec cette grille par trois évaluateurs différents et sur deux textes narratifs différents, suggère une bonne fiabilité de l'outil.
Mots-clés : prosodie en lecture, évaluation de la fluence, fluence en lecture

Echelle Multi-Dimensionnelle de Fluence : a new tool to assess reading fluency including prosody in french, calibrated for grade 2 to 5

Abstract

Children's reading fluency is usually assessed in classrooms using the instruction : "read as fast as you can". This instruction tends to perpetuate the confusion between fluency and speed. However, reading fast is not enough to be a good reader. Beyond speed, accuracy and automaticity, prosody, i.e. expressivity and phrasing, is a necessary skill to read fluently and understand the content. We present here an assessment tool for reading fluency that includes prosodic features, namely expressivity and phrasing. This tool is an adaptation of the MultiDimensionnal Fluency Scale (Rasinski, 2004) for French. Three raters assessed the reading of two narrative texts by 295 children from grade 2 to 7. Their scores are presented as a standardization of this scale. Various analyses of the scores, inter-rater agreement, correlation with classical standardized reading rate test and acoustic parameters, indicate a good fidelity, sensitivity and validity of this reading assessment tool.
Keyword : reading prosody, reading fluency, reading fluency assessment

Quand se pose la question de la définition d'un bon lecteur, la caractéristique revenant le plus fréquemment est que sa lecture "sonne" comme quelqu'un qui parle. Plus les paramètres de la lecture orale se rapprochent du discours parlé, plus le lecteur est perçu comme expert. Un lecteur expert est donc capable de reprendre les caractéristiques du langage parlé et de les utiliser dans sa lecture : vitesse, rythme, accentuation, variations de mélodie et d'intensité. Ces dimensions constituent ce que l'on appelle la prosodie du langage : la prosodie est l'interprétation orale du texte.

En classe, les compétences en lecture de l'enfant à partir du CE1 sont très fréquemment évaluées par des tests de "fluence", mesurant le nombre de mots correctement lus par minute à partir d'un texte étalonné, comme par exemple le test ELFE (Zorman *et al.*, 2008). Ainsi, l'entrainement à la fluence est assimilé à un entrainement à la vitesse. En recherche, la fluence est évaluée également par des mesures de vitesse, qu'elles soient utilisées pour caractériser la maitrise des deux voies de lecture, directe ou indirecte, en travaillant sur des listes de mots et de pseudomots (Gentaz *et al.*, 2013), ou pour caractériser une maitrise de la précision et de l'automaticité, en travaillant sur des textes (Fuchs *et al.*, 2001; Gentaz *et al.*, 2013; Kuhn, 2005). Cependant ce type d'évaluation et d'entrainement tend à entretenir la confusion entre fluence et vitesse, en particulier chez les plus jeunes. En effet, il n'est pas rare de rencontrer des enfants d'école élémentaire pour qui lire bien, c'est lire le plus vite possible, sans aucun souci de rythme, d'expressivité ou de compréhension. Or, la vitesse et la qualité du décodage sont une base nécessaire, mais non suffisante, au développement d'une lecture fluente. Les paramètres prosodiques tels que le rythme, le placement des pauses ou les variations d'intonation, d'intensité ou de volume, sont autant d'ingrédients tout aussi indispensables au développement de la lecture experte.

La bonne utilisation de la prosodie en lecture est en effet très utile, non seulement pour faire entendre une parole fluente à l'auditeur, mais aussi pour faciliter la compréhension du texte entendu. Par exemple, des pauses ou une intonation qui ne respectent pas la structure syntaxique d'un texte nuisent à cette compréhension. Le placement correct des pauses s'appuie énormément sur le repérage des ponctuations et l'intonation sur les ponctuations finales. Cependant, très souvent, cette ponctuation est insuffisante pour permettre une bonne gestion de la respiration, des pauses syntaxiques et de l'intonation. Il est alors nécessaire que le lecteur soit capable de découper le discours en unités syntaxiques, afin de produire despauses et une intonation respectant cette syntaxe. Cette gestion « en-ligne » de la respiration, des pauses et de l'intonation d'un texte demande de l'anticipation. C'est donc une opération cognitive complexe qui demande du temps et de l'entrainement pour se mettre en place (voir Godde *et al.* 2020a , pour une revue des études sur le développement de la prosodie en lecture).

Finalement la prosodie en lecture, i.e. la capacité à avoir une bonne intonation, et a fortiori de l'expressivité, est également très liée à la compréhension du texte par le lecteur. En effet, utiliser le ton adéquat, varier le volume et l'intensité de la voix de façon appropriée, suppose l'accès au sens du texte, car il y a peu d'indices typographiques de ces aspects du langage (Martin, 2011). Plusieurs études montrent un lien entre prosodie en lecture et compréhension écrite (Godde *et al.*, 2020a). Quelques études longitudinales semblent valider un lien bidirectionnel qui varie au cours du temps (Calet *et al.*, 2015b; Veenendaal *et al.*, 2016a). Durant les premières années d'apprentissage, ces études montrent peu de lien entre prosodie et compréhension, cette dernière étant alors plutôt liée aux capacités de décodage. A partir du CE2, ce lien se renforce. Il est alors plutôt dirigé de la compréhension vers la prosodie. Les enfants comprennent et adaptent leur rythme et leur prosodie au texte en fonction de leur compréhension. Puis, en fin de primaire, ce lien semble s'inverser et les jeunes lecteurs ont apparemment de plus en plus tendance à s'appuyer sur la prosodie pour s'aider à comprendre, particulièrement quand ils abordent des textes complexes (Benjamin et Schwanenflugel, 2010). Ce lien prosodie-compréhension se renforce ensuite dans le secondaire (Yildirim *et al.*, 2018).

Ainsi les jeunes lecteurs peuvent s'appuyer sur la prosodie pour renforcer leurs compétences en lecture, notamment la compréhension. Elle peut alors également aider l'auditeur à évaluer la compréhension du lecteur. Il est donc important de suivre l'acquisition de cette compétence dès le

début de l'apprentissage de la lecture, au même titre que sont suivies les acquisitions du décodage, de l'automaticité et de la compréhension. Les tests traditionnels, de type nombre de mots lus par minute, ne permettent pas d'évaluer la prosodie. Ils risquent même d'encourager la vitesse au détriment des autres compétences. Cependant, l'évaluation de la prosodie n'est pas une tâche facile. Dans ce texte, nous nous interrogerons sur les dimensions de la prosodie et les différents moyens de les évaluer, avant de proposer un nouvel outil d'évaluation de la fluence incluant la prosodie.

Évaluation de la prosodie

Pour suivre le développement de la prosodie, il convient d'adopter un outil d'évaluation adapté et suffisamment précis pour capter des changements dans les différents paramètres prosodiques. Deux questions se posent alors : quels paramètres évaluer pour avoir une bonne image des compétences prosodiques et comment évaluer ces paramètres.

Dimensions de la prosodie en lecture Les définitions récentes de la fluence prennent en compte la prosodie. Une des définitions les plus citées est la suivante : « Une composante critique de la lecture fluente est la capacité à lire de façon prosodique, c'est à dire avec une expressivité et une intonation appropriées, et un phrasé qui permet le maintien de la compréhension » (o.p. Kuhn *et al.* 2010, p.233, traduit de l'anglais). D'autres définitions intègrent les mêmes dimensions prosodiques : phrasé, intonation et expressivité (Dowhower, 1991; Rasinski, 2004).

Phrasé. Le découpage du discours en unités signifiantes est une fonction indispensable de la prosodie permettant la compréhension de l'auditeur. Ce découpage est principalement indicé par le placement et la durée des pauses ainsi que par l'allongement de certaines syllabes, notamment prépausales. Il est en général appelé phrasé. Le phrasé est donc intimement lié à la syntaxe du texte et à sa ponctuation.

On peut distinguer trois types de pauses : les pauses d'hésitations, les pauses respiratoires et les pauses syntaxiques. Les pauses d'hésitations sont généralement dues à des problèmes de décodage ou de compréhension, elles sont le signe d'une activité cognitive augmentée et leur nombre diminue au fur et à mesure du développement des compétences en lecture. Les pauses respiratoires sont des pauses indispensables. Elles sont plus nombreuses chez les jeunes enfants du fait d'une moindre capacité pulmonaire. Il est nécessaire pour les enfants d'apprendre à coordonner leur respiration avec la lecture et la syntaxe du texte. Cette acquisition est également progressive. Une étude de Grosjean et Collins (1979) montre l'effet néfaste d'une lecture trop rapide sur le placement des pauses respiratoires. Quand la lecture est trop rapide, le lecteur inspire quand ça lui est nécessaire, sans forcément respecter les frontières syntaxiques, ce qui entraine un phrasé erroné et nuit à la compréhension. Finalement, les pauses syntaxiques sont utilisées pour mettre en relief le sens du texte. Elles peuvent s'appuyer sur la ponctuation mais celle-ci est souvent insuffisante (FayoL, 1997; Martin, 2011). L'évaluation du phrasé doit donc permettre d'évaluer la capacité des enfants à découper le texte de façon licite et signifiante.

Intonation. L'intonation désigne les variations mélodiques de la voix. Cette intonation remplit de nombreuses fonctions. Elle permet notamment de marquer l'accentuation lexicale, indispensable en anglais, mais absente en français, considéré comme une langue avec un accent lexical final fixe (Encrevé, 1988; Dupoux *et al.*, 2001; Martin, 2009), les grandes frontières syntaxiques et la modalité

de la phrase, par exemple par la baisse de fréquence fondamentale en fin de phrase déclarative ou, au contraire, une montée en fin de phrase interrogative totale (i.e. non marquée par une inversion ou un mot interrogatif). Comme les pauses, l'intonation est très liée à la ponctuation et à la syntaxe du texte. Elle va donc également contribuer à marquer le phrasé. On parle alors de prosodie linguistique. Les variations mélodiques peuvent également permettre de mettre en avant (on parle d'emphase ou de focus étroit vs. large) ou en retrait une partie du texte (e.g. partie de texte marqué par des parenthèses ou tirets) mais aussi marquer des sentiments ou une intention du locuteur : sarcasme, doute, ironie, tristesse... Dans ce cas, ces variations d'intonation vont plutôt participer à l'expressivité, on parle alors de prosodie paralinguistique.

Expressivité. L'expressivité est certainement le plus subtil des paramètres prosodiques mentionnés dans les définitions présentées. Un lecteur peut tout à fait respecter le phrasé et l'intonation nécessaire au texte, sans aucune expressivité. Il sera alors compréhensible mais donnera l'impression d'une lecture plate et monotone. L'expressivité participe à la prosodie paralinguistique. Elle n'est pas liée directement à la syntaxe du texte mais plutôt à son sens. Son expression peut être indiquée par la ponctuation, par exemple les points d'exclamation, mais elle est surtout très liée à la compréhension du texte par le lecteur. En effet, bien souvent, il est nécessaire de comprendre l'histoire, voir d'inférer, par exemple, le contexte d'énonciation pour adopter l'expressivité adéquate. Seul le contexte peut ainsi permettre d'inférer le doute ou le sarcasme. Afin d'avoir une expressivité appropriée, qui donne vie au texte, le lecteur doit donc avoir une compréhension en profondeur du texte lu (Erekson, 2010). Il est important de noter que l'expressivité n'est pas figée pour un texte donné. Elle dépend du locuteur et de son interprétation du texte. En effet, il n'existe pas une seule façon de donner vie à un texte. Plusieurs lecteurs experts, proposant des figures d'expressivité très différentes, pourront être jugés également expressifs par les auditeurs. Il n'existe donc pas un but à atteindre, une règle d'or, mais plusieurs interprétations possibles d'un texte. C'est cette grande variabilité dans l'expressivité qui la rend plus difficile à appréhender par les jeunes lecteurs, et également plus complexe à évaluer.

Les définitions de la prosodie proposent donc une distinction entre ces trois dimensions prosodiques. Cependant le terme d'intonation réfère aussi bien à des éléments relevant du phrasé que de l'expressivité, considérés comme les deux dimensions fondamentales de la prosodie.

Comment évaluer ces dimensions ? Dans la littérature, les études s'intéressant au développement de la prosodie mettent en œuvre deux types d'évaluation : acoustique ou subjective. Le choix du type d'évaluation va dépendre de son objectif.

Analyses acoustiques. Les deux dimensions principales de la prosodie, le phrasé et l'expressivité, peuvent se traduire en termes de caractéristiques acoustiques : le phrasé par la durée et le placement des pauses, les variations d'intonation et l'expressivité par des variations de mélodie et d'intensité, etc. Cowie *et al.* (2002) ont analysé une quarantaine de marqueurs acoustiques de la prosodie, dans la lecture d'enfants de 8 à 10 ans, afin de voir lesquels étaient associés à une lecture expressive. Ils ont ainsi pu dégager deux grands groupes de marqueurs acoustiques :

- les marqueurs de fluence. Ce sont essentiellement des marqueurs rythmiques : durée et fréquence des pauses, taux de prononciation des syllabes, et fréquence des variations de fréquence fondamentale.
- les marqueurs d'expressivité. Ce sont essentiellement des marqueurs mélodiques liés aux variations de fréquence fondamentale : amplitude, durée, variations d'une phrase à l'autre.

Cowie *et al.* (2002) remarquent également que ces catégories sont interdépendantes : les paramètres temporels sont aussi, dans une moindre mesure, des marqueurs d'expressivité et inversement. En effet, l'automaticité est un pré-requis nécessaire à l'expressivité. Benjamin *et al.* (2013) ont également corrélé certains paramètres acoustiques à une évaluation subjective de la prosodie, chez des lecteurs de 7 à 9 ans. Ces corrélations font ressortir le même type de paramètres : les pauses marquent le phrasé et les variations de fréquence fondamentale, l'expressivité.

Dans la littérature, ce type d'analyse est particulièrement utilisé pour étudier le développement de la prosodie en lecture (e.g. Álvarez-Cañizo *et al.*, 2018; Schwanenflugel *et al.*, 2015). Il nécessite cependant l'enregistrement puis l'analyse acoustique de la lecture des enfants, donc du matériel adéquat et un temps d'analyse important. Cette méthode d'évaluation est donc plus particulièrement destinée à la recherche, et non à une utilisation dans les classes. Ces paramètres acoustiques peuvent cependant être utilisés pour le développement d'outils de visualisation de la prosodie, par exemple le logiciel WinPitch (Martin, 2005), ou d'évaluation automatique de la prosodie s'appuyant à la fois sur des critères objectifs et des échelles d'évaluations subjectives (Godde *et al.*, 2020a).

Échelle subjective. Les échelles d'évaluations subjectives sont des méthodes plus aisément utilisables pour évaluer la fluence dans le cadre scolaire et de la recherche. Plusieurs échelles d'évaluation de la fluence incluant la prosodie ont été proposées dans la littérature. Elles concernent principalement la langue anglaise mais peuvent être transposées dans d'autres langues. Certaines proposent des échelles unidimensionnelles (Pinnell *et al.*, 1995; Zutell et Rasinski, 1991). Elles ont en général quatre niveaux allant du lecteur débutant, lent et décodant difficilement, au lecteur expert, rapide et expressif. Elles ont été conçues pour permettre aux enseignants d'évaluer rapidement leurs élèves. Cependant elles sont assez imprécises car elles supposent que les enfants maîtrisent à un même niveau les différentes compétences de vitesse, de phrasé et d'expressivité, ce qui est rarement le cas. Afin d'avoir une évaluation plus précise du niveau des enfants, des chercheurs ont proposé des échelles multidimensionnelles (Benjamin *et al.*, 2013; Rasinski, 2004; Zutell et Rasinski, 1991). Ces échelles évaluent plusieurs paramètres permettant de donner une note globale tout en prenant en compte l'hétérogénéité des niveaux d'acquisition. Toutes ces échelles reprennent les dimensions de décodage, vitesse, phrasé et expressivité.

Ces échelles multidimensionnelles sont utilisées aussi bien dans les classes par des enseignants que par des chercheurs dans le cadre d'études portant sur le lien entre la prosodie et d'autres compétences de lecture (vitesse, compréhension...) (e.g. Veenendaal *et al.*, 2016b). Nous proposons ici d'adapter en français l'échelle « Multidimensionnal Fluency Scale », échelle multidimensionnelle créée pour l'anglais par Rasinski (2004). Nous analyserons les performances d'enfants du CE1 à la 5e, évalués avec cette échelle, afin de décrire les caractéristiques principales de ce nouvel outil.

Méthodologie

Participants Dans le cadre d'une étude sur le développement de la prosodie en lecture chez les enfants français, nous avons enregistré 323 enfants du CE1 à la 5e. Les caractéristiques de la cohorte : âges moyens, effectifs par niveau et nombre de filles et garçons, sont présentées dans le tableau 3.1. Les enfants ont été enregistrés dans leurs écoles : deux écoles primaires et un collège de l'agglomération grenobloise, situés dans une zone non-REP et accueillant des enfants de milieux socio-économiques diversifiés, avec l'autorisation de la directrice académique de l'Isère, des directrices d'écoles, du principal du collège et de tous les parents. Lors de cette étude nous avons uniquement conservé les enregistrements des enfants normo-lecteurs, sans diagnostic de dyslexie, de langue maternelle française ou bilingues dont l'une des langues maternelles est le français, soit 295 sujets.

TABLE 3.1 – Caractéristiques des participants en fonction de leur niveau scolaire – Participants characteristics according to their grade level

	CE1	CE2	CM1	CM2	6e	5e
nombre de sujets	61	47	54	63	33	37
filles	27	25	26	23	15	21
âge moyen (ET) (étendue)	7 :11(3,8) (7 :2-8 :5)	8 :11(3,4) (8 :5-9 :5)	9 :11(5,1) (8 :5-10 :5)	10 :11(4,5) (10 :3-12)	11 :11(4,8) (11 :4-13)	12 :8(5,4) (11 :8-13 :9)
vitesse de lecture	116(43)	160(62)	199(59)	237(84)	232(59)	287(76)
score PM38 (ET) (scores refs)	33.6(9.3) (33)	37.5(8) (36)	39.7(6.8) (40)	43.7(6.8) (41)	38.6(5.9) (38.9(6.4))	41.5(6.3) (42.01(7.1))

Moyenne (écart-type) de la vitesse de lecture (nombre de mots correctement lus en 2 min) et efficience cognitive générale (scores bruts moyens (écart-type) et en référence, le 50ème percentile pour le primaire et moyenne (écart-type) pour le collège, à l'étalonnage des matrices de RAVEN PM38)

Matériel

Prosodie Textes. Les deux textes utilisés pour évaluer la prosodie en lecture ont été écrits par les auteurs pour inciter à une lecture expressive, à partir d'une nouvelle de Bernard Friot (Histoires Pressées, 2007). Ils sont présentés en annexe A. Le texte 1 est un texte narratif de 174 mots. Il comporte treize phrases de sept à dix-huit mots. Le texte 2 est un dialogue de 168 mots entre deux personnages comportant dix questions et quatre exclamations. Il comporte vingt-trois phrases de deux à dix-sept mots. Ces deux textes ne présentent pas de difficultés lexicales particulières pouvant provoquer des difficultés de décodage. Chaque texte est présenté aux enfants sur une feuille A4 en police Arial, taille 14, interligne double et aligné à gauche.

Évaluation de la prosodie. Nous avons choisi d'adapter la « Multidimensionnal Fluency Scale » (MDFS), échelle multidimensionnelle créée pour l'anglais par Rasinski (2004). Cette échelle comporte quatre items : expressivité, phrasé, fluidité et rythme, notés de 1 à 4. La note 1 correspondant à une maitrise faible et la note 4 à une maitrise experte. Le score global maximal est donc de 16. Cette échelle a été choisie pour sa simplicité d'utilisation et le nombre limité d'items à évaluer, couvrant cependant l'ensemble des compétences attendues d'un jeune lecteur.

Dans sa version anglaise, cette échelle a été largement utilisée dans des études portant sur la relation entre la prosodie et les autres compétences de lecture (Godde *et al.*, 2020a) et traduite dans différentes langues (Álvarez-Cañizo *et al.* 2018; Veenendaal *et al.* 2016b; Yildirim *et al.* 2018, par exemple). Moser *et al.* (2014) ont évalué la fiabilité de cette échelle avec des CM1, pour différents types de texte. Ils ont pu montrer une forte cohérence entre évaluateurs, et une bonne fiabilité, quel que soit le type de textes lus. Ils ont également constaté une forte robustesse interne du test pour les textes narratifs (0.94 à 0.97) et informatifs (0.92 à 0.98). La MDFS est donc une échelle adaptée à l'évaluation individuelle. Cette étude préconise la lecture de 2 passages, évalués par plusieurs évaluateurs pour un rapport temps/fiabilité optimum.

Nous avons traduit et adapté l'échelle MDFS aux particularités de la langue française. Par exemple la notion d'accentuation lexicale, importante en anglais, mais absente en français, a été retirée des descriptions. La version définitive de cette échelle est présentée en Figure 3.1. Comme dans l'échelle anglaise, le score global maximum est de 16.

L'item expression évalue les variations de mélodie, d'intensité et de volume correspondant à l'interprétation du texte. L'item phrasé évalue le placement des pauses syntaxiques et respiratoires et l'intonation, dans leur respect de la syntaxe du texte. L'item décodage évalue la précision et l'automaticité de l'usage des correspondances grapho-phonémiques. Finalement, la vitesse évalue le nombre de mots lus par minute par rapport à un débit conversationnel.

Dans le cadre de notre étude, trois évaluateurs (tous enseignants) ont écouté les enregistrements des textes 1 et 2. Ils ont évalué la fluence des lecteurs sur chacun des enregistrements à l'aide de l'échelle multidimensionnelle de fluence (EMDF) présentée en Figure 3.1. Chaque enregistrement est écouté pendant une minute et une note de 1 à 4 est attribuée à chaque item de l'échelle. L'évaluation se déroule en deux phases. Lors de la première phase, les évaluateurs écoutent une dizaine d'enregistrements sur lesquels ils discutent afin de s'entendre sur les scores attribués à chaque item. A l'issue de cette phase d'accord, les trois évaluateurs attribuent, indépendamment les uns des autres, un score à chaque enregistrement.

Dimension	1	2	3	4
Expression	Pas d'expression et peu d'enthousiasme dans la voix. Lit les mots comme ils arrivent. N'essaye pas de faire sonner comme du langage parlé. Monotone.	Un peu d'expression. Commence à utiliser sa voix pour sonner naturel dans certaines parties du texte mais pas partout.	Passage lu en partie avec expression. Sonne comme du langage parlé pour la plupart du texte. Volume de la voix approprié.	Lit avec une bonne expression et de l'enthousiasme durant tout le texte. Sonne comme du langage parlé.
Phrasé	Mot à mot fréquent. Intonation monotone.	Fréquents phrasés de 2 à 3 mots, une lecture hachée. Intonation inappropriée qui ne marque pas la fin des phrases et des propositions.	Mélange de longs phrasés sans pause et d'arrêts inappropriés pour respirer. Nombreuses pauses ne respectant pas la syntaxe. Intonation en partie adaptée.	Phrasé généralement correct, respectant la syntaxe du texte. Intonation adaptée.
Décodage	Longues pauses fréquentes, hésitations, faux départs, chuchotements, répétitions, essais multiples, mots modifiés.	Plusieurs points difficiles dans le texte avec hésitations, longues pauses, erreurs de lecture.	Rupture occasionnelle de la fluidité sur des mots et/ou des structures particulières. Quelques erreurs de lecture.	Lecture généralement fluide avec quelques ruptures mais résolues rapidement, en général en s'auto-corrigeant.
Vitesse	Laborieux.	Lent.	Mélange irrégulier de lecture lente et rapide.	Constant et Conversationnel.

FIGURE 3.1 – Echelle Multi-Dimensionnelle de Fluence (EMDF), adaptation de l'échelle MultiDimensionnelle de Fluence (MDFS) de Rasinski (2004) pour le français - Echelle Multi-Dimensionnelle de Fluence (EMDF), adapted from the MultiDimensionnal Fluency Scale (MDFS) Rasinski (2004) for French.

Tests contrôles **Vitesse de lecture.** Le texte utilisé pour évaluer la vitesse de lecture, nombre de mots correctement lus (NMCL), est le texte « La Mouette » tiré de la batterie d'évaluation des troubles du langage Evaléo 6-15 (Maeder *et al.*, 2018). « La Mouette » est un texte signifiant de 450 mots, composé de trois paragraphes. Il est demandé aux enfants de lire pendant deux minutes en essayant d'aller le plus loin possible dans le texte. L'examinateur relève les erreurs et omissions pendant la lecture, ainsi que le nombre de mots lus à l'issue des deux minutes. Il peut alors calculer le NMCL en deux minutes. Les résultats de vitesse de lecture sont présentés dans le tableau 3.1. Ils sont conformes aux scores moyens des niveaux scolaires des enfants.

Efficience cognitive générale. L'efficience cognitive générale a été mesurée avec le test des matrices de Raven PM38 (Raven *et al.*, 2003). Ce test a été réalisé en passation collective, sous la supervision de l'examinateur et de l'enseignant. Les résultats de ce test sont présentés dans le tableau 3.1. Ils sont globalement conformes à leur classe d'âge. Il est cependant nécessaire de noter que le score brut moyen des CM2 est supérieur au 50éme percentile donné par l'étalonnage (Raven *et al.*, 2003).

Procédure Les sujets ont été enregistrés individuellement dans un lieu calme à proximité de leur classe. Durant les lectures, la voix des enfants a été enregistrée avec un micro Schure Beta 53, un amplificateur Berhinger MIC100, relié à un Biopac MP150 permettant d'enregistrer simultanément leur respiration (pour l'étude, non développée dans le présent article, des pauses respiratoires) grâce à deux ceintures abdominales et thoraciques qui ne gênent en rien la lecture à voix haute. Lors de cet enregistrement, l'expérimentateur leur demandait de lire les textes 1 et 2, dans un ordre aléatoire. La consigne donnée aux enfants était de lire comme s'ils lisaient une histoire à des élèves de maternelle. Il était ensuite demandé aux enfants de lire le texte permettant d'évaluer leur vitesse de lecture. La consigne était de lire pendant deux minutes en allant le plus loin possible dans le texte.

Résultats

Résultats globaux à l'EMDF en fonction du niveau scolaire L'accord entre les scores donnés par les trois évaluateurs à chaque lecture est calculé avec l'alpha de Krippendorf pour données ordinales (Hayes et Krippendorff, 2007). Ce coefficient se lit comme une corrélation : plus la valeur est proche de 1, meilleur est l'accord. Les valeurs de ce coefficient pour chacun des paramètres sont renseignées dans le tableau 3.2.

Pour le texte 1, on peut remarquer que la majorité des coefficients est supérieure à 0.8, ce qui indique un très bon accord inter-juges (Hayes et Krippendorff, 2007). Pour le texte 2, les coefficients pour chaque paramètre sont plus faibles (0.7) que pour le score total (0.9). En effet, quand les scores des évaluateurs divergent, c'est, dans la majorité des cas, d'un point sur un ou deux items Mais ces désaccords se compensent pour donner un score final sur 16 identique, entrainant un très bon accord inter-juge sur le score total. Pour conclure, le score total sur 16 constitue, pour les deux textes, une mesure fidèle de la prosodie des enfants.

Les tableaux 3.3 et 3.4 présentent les scores de fluence moyens et quantiles, obtenus avec l'échelle proposée, par niveau scolaire pour chacun des textes.

Le tableau 3.5 présente les coefficients de corrélation entre les résultats aux deux textes. La distribution des scores par niveau est présentée en Figure 3.2. On peut voir sur ces distributions

TABLE 3.2 – Accord inter-évaluateur pour chaque paramètre évalué et pour chaque texte – Inter-rater agreement for each parameters and each text.

	Texte 1	**Texte 2**
Expression	0.82	0.71
Phrasé	0.80	0.72
Décodage	0.78	0.71
Vitesse	0.81	0.75
Total	0.95	0.93

Coefficient alpha de Krippendorf pour données ordinales (Hayes et Krippendorff, 2007)

TABLE 3.3 – Moyenne et quantiles des scores de fluence EMDF sur le texte 1 par niveau – EMDF fluency scores's averagesand quantiles for text 1 according to grade level

Texte 1	CE1	CE2	CM1	CM2	6e	5e
90%	11.3	14.1	13.7	15.0	14.7	16.0
75%	10.0	12.6	12.6	14.0	14.2	15.1
50%	7.6	11.3	11.6	12.6	13.0	14.0
25%	6.1	9.0	11.0	11.0	11.8	12.6
10%	5.3	6.6	9.5	9.9	11.0	11.2
moyenne	8.1	10.6	11.6	12.3	12.7	13.7
écart-type	2.3	2.9	1.8	2.4	1.8	1.7

TABLE 3.4 – Moyenne et quantiles des scores de fluence EMDF sur le texte 2 par niveau – EMDF fluency scores's average and quantiles for text 2 according to grade level

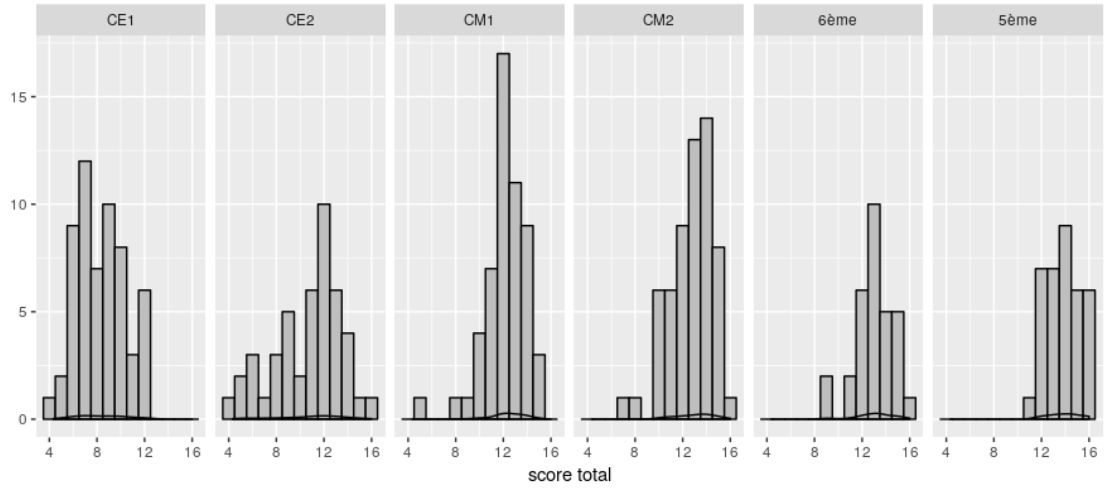
Texte 2	CE1	CE2	CM1	CM2	6e	5e
90%	11.7	13.9	15.0	15.1	15.3	16.0
75%	10.3	13.0	14.0	14.3	14.7	15.1
50%	8.7	11.6	13.3	13.7	13.3	14.0
25%	7.3	9.7	12.3	12.2	12.7	13.3
10%	6.0	6.1	10.8	11.6	11.7	12.7
moyenne	8.8	10.9	12.9	13.3	13.4	14.1
écart-type	2.2	2.9	1.9	1.6	1.6	1.3

Table 3.5 – Coefficient de corrélation et corrélation partielle entre les scores EMDF aux textes 1 et 2 – Correlation and partial correlation between text 1 and 2

	expressivité	phrasé	décodage	vitesse	total
corrélation	0.70***	0.79***	0.72***	0.76***	0.83***
corrélation partielle	0.65***	0.68***	0.54***	0.61***	0.74***

Coefficient de corrélation (Spearman) et coefficient de corrélation partielle (Spearman), prenant en compte le niveau scolaire. Seuil de significativité : * $p < .05$, ** $p < .01$, *** $p < .005$

Figure 3.2 – Distribution des scores totaux moyens à l'EMDF pour chaque niveau – Distribution of average EMDF total scores for each grade level.



que l'échelle montre une bonne sensibilité pour les plus jeunes, du CE1 au CM1, mais que celle-ci tend à diminuer à partir du CM2. On constate un effet de saturation : les scores des élèves étant de plus en plus élevés, la distribution se déplace vers les scores élevés et discrimine de moins en moins bien les bons lecteurs, mais garde une bonne sensibilité pour les plus faibles lecteurs. On peut également remarquer que les scores sont plus élevés pour le texte 2. Cela peut s'expliquer par le fait que ce texte est un dialogue. Sa mise en page, avec retour à la ligne régulier, facilite un phrasé correct. Le dialogue incite également les enfants à plus d'expressivité. On s'attend donc à des scores plus élevés en phrasé et en expressivité sur ce type de texte.

On peut constater dans le tableau 3.5 que la corrélation des scores entre les 2 textes est plutôt bonne (supérieure à 0.7). Les items les moins corrélés sont l'expressivité et le décodage. En effet, ce sont deux items très dépendants du texte. Le texte 2, présenté en annexe A, est un dialogue, appelant ainsi plus d'expressivité que le texte 1, uniquement narratif. Les coefficients de corrélation partielle, prenant en compte le niveau scolaire des lecteurs, sont plus faibles. Afin de fournir une indication du niveau de fluence des élèves en fonction du niveau scolaire, qui soit plus générale et moins dépendante du type de texte, nous présentons dans le tableau 3.6 les scores moyens sur les deux textes.

Relation entre les scores à l'EMDF et la mesure classique de vitesse de lecture Le tableau 3.7 présente les corrélations entre le score moyen des quatre différentes dimensions de l'EMDF (ainsi que le score global) et le score de vitesse de lecture, mesuré avec le texte « La

Table 3.6 – Moyenne et quantiles des scores de fluence EMDF sur la moyenne des 2 textes par niveau – EMDF fluency scores for both text , quantiles and average, according to grade level

	CE1	CE2	CM1	CM2	6e	5e
90%	11.4	14	14	15	15	15.9
75%	10	12.7	13.5	14.1	14	14.8
50%	8.4	11.5	12.4	13.3	13.2	13.9
25%	6.9	9	11.5	11.8	12.2	12.9
10%	5.8	6	10	10.3	11.3	12
moyenne	8.4	10.8	12.3	12.8	13.1	13.9
écart-type	2.1	2.8	1.8	1.8	1.6	1.3

Table 3.7 – Coefficients de corrélation entre score de fluence Evaleo 6-15 et de fluence EMDF – Correlations between Evaleo 6-15 and EDMF fluency scores

	CE1	CE2	CM1	CM2	6e	5e
expressivité	0.68***	0.77***	0.23	0.44***	0.26	0.50**
phrasé	0.81***	0.83***	0.45***	0.63***	0.44*	0.55***
décodage	0.74***	0.82***	0.56***	0.70***	0.69***	0.75***
vitesse	0.84***	0.87***	0.52***	0.71***	0.69***	0.47***
total	0.86***	0.88***	0.49***	0.72***	0.53***	0.65***

Coefficients de corrélation (Spearman). Seuil de significativité : * $p < .05$, ** $p < .01$, *** $p < .005$

Mouette » (Maeder *et al.*, 2018) qui mesure un nombre de mots correctement lus, mais n'inclut aucun paramètre de prosodie.

Les scores globaux avec et sans prosodie sont fortement corrélés ($r > 0.8$) en CE1 et CE2. Par contre, cette corrélation diminue quand le niveau augmente ($r < 0.7$), ce qui suggère qu'avec l'augmentation du niveau de lecture, les dimensions prosodiques deviennent de plus en plus indépendantes de la vitesse de lecture. Si on s'intéresse à chaque paramètre de l'échelle, on retrouve ce même schéma. Le décodage et la vitesse sont les paramètres les plus corrélés au score de fluence à tous les niveaux, et ceci de façon relativement stable. Cela confirme que la fluence, au sens du nombre de mots correctement lus par minute, est bien captée par cette échelle. La corrélation avec le phrasé est également élevée en CE1 et CE2 ($r > 0.8$) mais diminue à partir du CM1 ($r < 0.6$). Cela peut s'expliquer par les hésitations plus nombreuses en CE1 et CE2, perturbant un phrasé correct et diminuant la vitesse de lecture. L'expressivité est le paramètre le moins corrélé à la vitesse de lecture. En CE1 et CE2, la corrélation est encore significative mais reste plus faible (autour de 0.7) que pour les autres paramètres (supérieur à 0.8). Cependant à partir du CM1, cette corrélation diminue fortement, devenant faible ($r < 0.5$), voire inexistante. Ces scores nous montrent donc que les paramètres d'expressivité et de phrasé apportent une information supplémentaire sur les compétences en lecture, par rapport à un score de nombre de mots correctement lus par minute. En CE1 et CE2, la prosodie en lecture est liée aux performances en vitesse de lecture, mais c'est de moins en moins le cas lorsque le niveau en lecture augmente.

Relation entre les scores à l'EMDF et les paramètres acoustiques Dans le cadre d'une étude longitudinale sur le développement de la prosodie (Godde *et al.*, 2020a), nous avons exploré

les liens entre les scores de chaque dimension de l'échelle multidimensionnelle et les paramètres acoustiques caractérisant la prosodie. Nous avons effectué des régressions logistiques ordinales pour déterminer quels paramètres acoustiques et verbaux mesurés pendant la lecture permettent de rendre compte des scores subjectifs obtenus par l'enfant aux différentes dimensions de l'EMDF. Les paramètres acoustiques inclus dans les régressions étaient les paramètres rythmiques et mélodiques. Les paramètres verbaux étaient le NMCL par minute (NMCLM), le nombre de répétitions et le nombre d'erreurs par minute. Ces régressions montrent un effet principal du NMCLM ($p < .001$) sur le jugement du score de vitesse et un effet principal du nombre d'erreur et de répétition sur le jugement du score de décodage ($p < .001$). Ces deux dimensions de l'échelle (vitesse et décodage) sont donc bien liées respectivement à l'automaticité et à la précision en lecture. Les coefficients de corrélation de Spearmann entre les modèles obtenus et nos données expérimentales sont élevés (.85 et .83). Pour le score de jugement de phrasé, on observe un effet principal des paramètres acoustiques rythmiques ($p < .001$) avec une corrélation entre scores modélisés et données expérimentales de .86. Finalement, l'analyse montre un effet principal des paramètres acoustiques mélodiques sur le score d'expressivité ($p < .001$) avec un coefficient de corrélation de .86. Ces deux dernières dimensions de l'échelle (phrasé et expressivité) sont donc bien liées aux paramètres acoustiques caractéristiques de la prosodie, respectivement les paramètres rythmiques et mélodiques. Les différentes dimensions de l'échelle sont donc bien ancrées dans un schéma acoustique et objectif correspondant à leurs définitions.

Discussion

L'échelle de mesure de la fluence présentée ici repose sur l'évaluation subjective de quatre paramètres. Les deux premiers paramètres, vitesse et décodage, sont des paramètres traditionnels des mesures de fluence. Mesurés quantitativement par le nombre de mots lus correctement en une minute, nous en proposons ici une évaluation qualitative. Les deux autres paramètres, expressivité et phrasé, apportent des informations supplémentaires sur les compétences en lecture des enfants évalués. En effet, si les scores obtenus à cette échelle sont très liés aux scores traditionnels de fluence (i.e. nombre de mots lus par minute) dans les premières années, cette corrélation diminue avec le développement des compétences en lecture. Le nombre de mots lus par minute n'explique alors qu'une part de ces compétences. L'EMDF, en prenant en compte les compétences prosodiques, permet d'accéder à une évaluation plus large de la lecture que la simple mesure de vitesse. Une lecture rapide n'est pas forcément une lecture qualitative et le score du nombre de mots lus par minute n'est pas toujours révélateur des qualités de lecteur des enfants. Évaluer séparément ces quatre paramètres permet également de mettre en lumière des stratégies différentes : des lecteurs qui lisent lentement (score vitesse faible) mais avec précision (décodage élevé) et avec intonation et des lecteurs qui privilégient la vitesse (vitesse élevée) au détriment de la précision de lecture (décodage faible), du phrasé et de l'intonation. La seule mesure du nombre de mot lus par minute ne permet pas cette analyse qualitative plus fine. Nous pensons donc que cette échelle est un outil intéressant comme complément au nombre de mot lus correctement par minute. Par ailleurs, les compétences prosodiques sont d'autant plus pertinentes à évaluer qu'elles sont liées à la compréhension écrite et que ce lien s'accentue avec le développement des compétences en lecture (Benjamin et Schwanenflugel, 2010; Yildirim *et al.*, 2018). Il est donc important de suivre l'acquisition de ces compétences.

Sensibilité, fidélité, validité Cette échelle présente une bonne sensibilité du CE1 au CM1, avec une répartition des scores et une moyenne permettant une bonne discrimination des faibles

jusqu'aux très bons lecteurs. Du CM2 à la 5e, la sensibilité diminue pour les bons lecteurs car les textes proposés, probablement un peu trop faciles pour des élèves de cet âge, ne permettent plus de les discriminer avec précision. Cependant la sensibilité reste bonne pour les lecteurs les plus faibles, ce qui peut en faire un bon outil de repérage des lecteurs en difficulté jusqu'en 5e.

Cette échelle présente également une bonne fidélité, avec un accord inter-juge très élevé, particulièrement sur le score total, et une bonne corrélation entre les deux textes évalués. Ces deux textes peuvent donc être utilisés séparément, comme deux tests différents, ou ensemble pour donner un score moyen plus robuste.

Finalement, l'échelle proposée ici est traduite d'une échelle construite originellement pour l'anglais. Chaque item évalué correspond à des dimensions de fluence (vitesse et décodage) et de prosodie (expressivité et phrasé). Cette échelle a donc une validité de contenu par sa construction. La corrélation des résultats aux différentes dimensions de l'échelle avec les mesures de fluence sans prosodie, sur le texte « La Mouette », montre une validité empirique. En effet, les dimensions de vitesse et décodage sont très corrélées au score de « La Mouette ». Les régressions montrent également que ces paramètres sont principalement liés au nombre de mots lus par minute et au nombre d'erreurs dans la lecture des textes proposés, alors que les scores d'expressivité et de phrasé, moins corrélés à ces scores de vitesse et décodage (particulièrement chez les plus grands) mais liés à des paramètres rythmiques et mélodiques, apportent bien des informations supplémentaires sur les compétences en lecture orale. Ce lien entre les scores obtenus avec cette échelle et les paramètres acoustiques et de mesure de vitesse et de précision sur les textes évalués attestent également une validité théorique certaine de cette échelle.

Utilisation L'échelle multidimensionnelle de fluence présentée ici nécessite un protocole d'utilisation particulier. Plusieurs points sont à contrôler pour obtenir un score représentatif des compétences en lecture du sujet.

L'évaluateur doit veiller aux points suivants :

- L'évaluation porte sur la première minute de lecture. En effet, la fluence, et en particulier la qualité de la prosodie, tend à diminuer au fil de la lecture quand le texte est long ou complexe, et ce d'autant plus que le lecteur est en difficulté (Benjamin et Schwanenflugel, 2010).
- Une note de 1 à 4 est attribuée à chaque item en fonction de la grille donnée en figure 3.1. Il sera préférable de juger du phrasé et de l'expressivité sur une partie de cette première minute présentant peu de difficulté de décodage. Le décodage est jugé sur la partie la plus complexe de cette minute de lecture.
- Cette échelle peut être utilisée par un seul évaluateur, par exemple pour le suivi d'élèves. Cependant, dans le cadre de la recherche, ce type d'échelle est, en général, utilisé avec plusieurs évaluateurs. Haskins et Aleccia (2014) pointent la difficulté d'une évaluation fiable avec les échelles d'évaluations subjectives, particulièrement quand il n'y a pas d'entrainement des évaluateurs. Les préconisations pour avoir des résultats aussi fiables que possible sont la présence de trois évaluateurs, préalablement entrainés, et une évaluation portant sur deux textes différents (Haskins et Aleccia, 2014; Moser *et al.*, 2014). Pour entrainer les évaluateurs, on présente une dizaine de lectures, couvrant des niveaux différents, sur lesquelles ils s'accordent pour attribuer un score à chaque item. A l'issue de cette phase d'accord, les évaluateurs notent chaque lecture indépendamment les uns des autres.

Limites d'utilisation L'échelle multidimensionnelle de fluence incluant la prosodie proposée ici est donc un outil d'évaluation apportant de nouvelles informations sur les compétences en lecture.

Comme vu dans un paragraphe précédent, la sensibilité de cette échelle en fait un bon outil d'évaluation pour les plus jeunes, du CE1 au CM1. Au-delà, sa capacité de discrimination est plus faible pour les bons lecteurs, elle sera alors plus adaptée pour un repérage des élèves en difficulté. Il serait intéressant de tester cette échelle avec des textes plus complexes, adaptés au niveau de lecture des collégiens. Il faut également noter que l'étalonnage porte sur un effectif limité, particulièrement pour le collège. La récolte de donnée avec l'EDMF à une plus grande échelle permettra d'obtenir un étalonnage plus précis sur les textes proposés.

Dans le cadre d'un suivi, donc de mesures répétées, les évaluations devront être réalisées à un intervalle permettant des progrès significatifs. En effet, les scores sont des scores discrets sur 4 points. Ils permettent donc de mesurer des progrès suffisants pour passer d'une catégorie à une autre, mais ne sont pas conçus pour évaluation fine des compétences. De plus, si les mesures répétées sont réalisées par un seul juge, par exemple un enseignant pour le suivi d'un élève, il est nécessaire que ce même juge effectue toutes les évaluations. Si le même juge ne peut évaluer toutes les mesures répétées, il est alors préférable que l'enfant soit évalué sur deux lectures par deux ou trois juges pour aboutir à une mesure fidèle permettant le suivi de l'évolution des compétences (Moser *et al.*, 2014). En effet, la subjectivité de l'évaluation, particulièrement de l'expressivité, ne permet pas un changement de juge aisé dans le cadre d'un suivi.

Conclusions

Nous proposons ici une échelle d'évaluation de la fluence prenant en compte la prosodie. L'utilisation de cette échelle permet d'évaluer des compétences en lecture, expressivité et phrasé, qui vont au-delà des simples indicateurs de vitesse et précision. Ces compétences prosodiques apportent notamment des informations sur les capacités de compréhension et d'anticipation des jeunes lecteurs.

Elle offre donc une vue élargie des compétences, comme des difficultés potentielles des élèves. Sa sensibilité, sa fidélité et sa validité en font aussi bien un outil utile en classe pour l'évaluation et le suivi des élèves qu'un outil pour la recherche. Cette adaptation permet d'utiliser, en français, un outil largement utilisé en anglais, notamment pour des études longitudinales.

Remerciements

Ce travail a été effectué dans le cadre du projet e-FRAN Fluence financé par la Caisse de dépôts et Consignations (CDC). Nous remercions les éducateurs et élèves pour l'accueil réservé à Erika Godde et Manon Metz, expérimentatrices. Nous remercions Manon Metz pour son aide dans les passations, Anne-Claire Dugué et Cécile Lebourg pour leur aide dans les évaluations subjectives.

TABLE 3.8 – Accord interévaluateurs (coefficient alpha de Krippendorf) pour chaque type de jury et chaque dimension de l'EMDF

jury	juges	lectures	expressivité	phrasé	décodage	vitesse	total
Experts 1	3	58	.73	.64	.54	.76	.90
Experts 2	3	61	.51	.75	.41	.53	.70
Enseignantes	11	51	.66	.58	.55	.72	.80
Etudiantes	9	181	.60	.60	.48	.58	.72

3.1.2 Différents types d'évaluateurs

Dans l'article précédent, les lectures avaient été jugées par trois évaluateurs "experts" en lecture d'enfants, enseignants et participant au projet et donc très familiers avec l'échelle et ses différentes dimensions. Le choix des juges pourrait donc expliquer des coefficients d'accord inter-évaluateurs élevés. Nous avons donc voulu tester l'utilisation de l'EMDF avec des juges de différents horizons. Cela nous permet de comparer les coefficients d'accord entre différents types de juges, les scores moyens attribués et le classement des lectures.

Méthodologie

Cette étude a été menée sur le texte 1 uniquement, avec des lectures issues du corpus longitudinal. Deux sessions d'évaluations supplémentaires ont été menées pour ce texte. Lors de la première session, neuf étudiantes en master 2 neuropsychologie de l'enfant ont évalué les 181 lectures du corpus longitudinal. Lors d'une deuxième session, onze enseignantes du premier degré, ont évalué 51 lectures de ce même corpus. Nous disposons également des évaluations de trois juges experts pour les deux premières années du corpus longitudinal. Pour chacune de ces sessions, le protocole a été le même. La session commence par une présentation et explication de l'utilisation de l'EMDF. Les juges écoutent ensuite collectivement 15 lectures d'entraînement. A la suite de chaque lecture, ils échangent sur les scores attribués pour se mettre d'accord sur la façon de noter. Il est intéressant de noter que pour le jury d'enseignantes, les scores ont rapidement été harmonieux entre les juges. A la suite de ces 15 lectures d'entrainement, l'écoute des lectures est toujours collective mais les notations sont alors individuelles, sans aucun échange entre les juges. Pour ces deux sessions les lectures du corpus longitudinal sont présentées dans un ordre aléatoire, de façon à ce que le même sujet ne soit pas entendu plusieurs fois à la suite et à ce qu'il n'y ait pas de progression en terme d'âge ou de niveau de lecture.

Résultats

Pour chaque jury et pour chaque année du corpus longitudinal, nous avons calculé l'accord inter-évaluateur en utilisant le coefficient Alpha de Krippendorf (Hayes et Krippendorff, 2007). Le tableau 3.8 présente ces coefficients pour chaque jury. Le jury expert étant composé de personnes différentes suivant les années, il a été scindé en deux. Le jury experts 1 est le jury ayant évalué les lectures du corpus transversal, ce qui correspond à l'année 1 du corpus longitudinal, soit uniquement des CE1. Le jury experts 2, ayant deux personnes en commun avec le jury experts 1, a évalué les lectures de l'année 2 du corpus longitudinal, soit uniquement des CE2. Pour rappel, le corpus longitudinal comprend trois lectures de 62 enfants enregistrés en CE1, CE2 et CM1.

Table 3.9 – Scores moyens à l'EDMF pour chaque jury et chaque dimension de l'EMDF

jury	expressivité	phrasé	décodage	vitesse	total
Experts	2(0.79)	2.47(0.76)	2.63(0.72)	2.61(0.69)	9.79(2.36)
Enseignantes	2.16(0.83)	2.59(0.69)	3.04(0.67)	2.69(0.83)	10.48(2.69)
Etudiantes	2.65(0.8)	2.91(0.69)	3.13(0.66)	3.07(0.84)	11.76(2.76)

Tout d'abord, concernant les coefficients d'accord inter-évaluateurs, on peut observer une différence entre les deux jurys d'experts. Celle-ci pourrait s'expliquer par un changement dans le protocole d'évaluation. En effet, lors du jury experts 2, la phase d'entrainement avait été écourtée, car l'accord était rapide entre les évaluateurs et ceux-ci étaient habitués à utiliser l'EMDF. D'autre part, si une partie des évaluations avait été réalisée collectivement à la suite de l'entrainement, une autre partie des évaluations a été réalisée quelques jours plus tard individuellement par un des évaluateurs. Cet "assouplissement" dans le protocole d'évaluation pourrait expliquer la différence entre les deux jurys experts. Une autre explication possible à cette différence est que le jury experts 1 a évalué les lectures des CE1 parmi des lectures d'enfants allant du CE1 à la 5^{e}. La variabilité dans la fluence des élèves était alors très importante, et les CE1 représentant les plus faibles lecteurs, il y a pu y avoir un effet plancher. Le jury expert 2 n'a évalué que des lectures de CE2, présentant donc une variabilité moindre. Plus globalement, on peut constater que les experts 1 ont des coefficients plus élevés que les enseignantes et les étudiantes. Le coefficient indique un bon accord sur le score total (>.70) mais cet accord est moyen, voir faible (< .60) pour les dimensions considérées indépendamment les unes des autres. Sans considérer les experts 2, dont on a évoqué les conditions particulières d'évaluation, on peut voir que la dimension présentant l'accord le plus faible est toujours le décodage (< .60), puis le phrasé et l'expressivité et la vitesse. Si l'accord est un peu plus faible pour les étudiantes, il reste assez proche pour les trois groupes d'évaluateurs.

Nous avons ensuite comparé les scores moyens donnés par chaque type de jury pour les mêmes 33 lectures du texte 1 évaluées par tous les jurys. Ces résultats sont présentés dans le tableau 3.9. L'effet du jury sur le score a été confirmé en comparant deux modèles de prédiction du score : un modèle unique et un modèle comprenant l'effet fixe du jury. Ces modèles ont été estimés avec la fonction lmer (package lme4 dans R) et les sujets comme paramètres aléatoires. L'effet du jury est significatif pour le score global ($\Delta\chi^2(2) = 28.2$, $p < .001$) ainsi que pour chacune des dimensions : expressivité ($\Delta\chi^2(2) = 38.6$, $p < .001$), phrasé ($\Delta\chi^2(2) = 15.4$, $p < .001$), décodage ($\Delta\chi^2(2) = 17.5$, $p < .001$) et vitesse ($\Delta\chi^2(2) = 17$, $p < .001$). A l'exception du décodage, les étudiantes donnent des scores plus élevés que les enseignants et les experts ($p < .001$ pour toutes les dimensions). Il n'y a pas de différence significative entre les scores des experts et des enseignants. Pour le décodage, les experts donnent des scores significativement plus faibles ($p < .001$) que les étudiantes et les enseignantes, entre lesquelles il n'y a pas de différence significative.

Afin d'aller plus loin dans la comparaison entre les différents types d'évaluateurs, nous nous sommes interressés à la corrélation entre les moyennes obtenues par les lecteurs pour chaque jury. Le tableau 3.10 présente les coefficients de Spearmann calculés en corrélant les scores des jurys deux à deux.

Il apparait dans ce tableau que les corrélations sur les scores globaux moyens sont très bonnes quels ques soient les jurys comparés (de l'ordre de .90). Ils sont également bons sur chaque dimension (compris entre .77 et .91), à l'exception du phrasé et du décodage entre les enseignantes et les experts, un peu plus faible. On voit donc que si les scores différent entre les types d'évaluateurs, tous classent les lecteurs de la même façon, particulièrement sur le score global.

Table 3.10 – Corrélation (coefficient de Spearmann) entre les scores moyens à l'EDMF sur le corpus longitudinal, pour chaque type de jury, deux à deux.

	nombre sujets	expressivité	phrasé	décodage	vitesse	total
Ens/Exp	34	.77***	.69***	.61***	.78***	.89***
Ens/Etud	51	.91***	.85***	.88***	.89***	.94***
Exp/Etud	119	.78***	.84***	.78***	.84***	.92***

Note : Ens = enseignantes, Exp = experts, etud = étudiantes. *** : $p < .001$

Discussion

Ces analyses supplémentaires, en jouant sur la connaissance des lectures d'enfants d'une part et de la prosodie et fluence en lecture d'autre part, nous permettent de confirmer la fidélité de l'outil présenté ici. En effet, on peut constater qu'avec un entrainement adapté, l'accord inter-évaluateurs est bon, quelle que soit l'expérience préalable des juges intervenant dans l'évaluation. Comme constaté dans l'article précédent sur le corpus transversal, cet accord est meilleur pour le score global. En effet, on observe le même effet pour tous les types de jurys, les scores ne différant très souvent que d'un seul point sur certaines dimensions et ces points de différence se compensant pour donner le même score global. Les étudiantes, jury le moins familier dans l'évaluation des lectures d'enfants, a tendance à donner des scores plus élevés que les jurys plus familiers de cet exercice. Malgré cela, les corrélations entre les scores moyens donnés par chaque type de jury sont élevés. Ainsi même si les scores différent d'un demi-point sur les dimensions de l'EMDF, le classement des enfants est le même pour tous les jurys. Ainsi l'EMDF peut-être efficacement utilisée par différents types d'évaluateurs. Cette étude complémentaire sur le type d'évaluateurs vient en appui des résultats obtenus pour le corpus transversal.

Cette étude met également en avant deux points importants. Tout d'abord le faible accord entre les deux groupes d'experts montre l'importance de l'entrainement et du suivi du protocole dans la mise en œuvre de l'outil. Le faible accord entre évaluateurs sur chacune des dimensions prises séparément, montre également l'importance, notamment pour la recherche, d'avoir plusieurs évaluateurs et de considérer un score moyen pour chacune de ces dimensions. Ces observations confirment donc les préconisations faites par Moser *et al.* (2014) sur la nécessité que plusieurs juges évaluent plusieurs textes pour chaque sujet. D'un point de vue plus large, cette étude met aussi en évidence la difficulté de mesures répétées, par exemple dans le cadre d'un suivi. En effet, on a pu constater qu'une mesure ponctuelle peut-être fiable quel que soit le jury et peut permettre de comparer les résultats de plusieurs sujets. Cependant pour mesurer les progrès dans le temps d'un même sujet, il risque d'être complexe de différencier la variabilité due au jury de celle due aux progrès de l'enfant. Nous reviendrons sur cette question dans le chapitre suivant.

3.1.3 Développement de la prosodie

Dans l'article présenté en début de chapitre, la fluence et la prosodie n'ont été abordées que du point de vue de l'étalonnage des textes et pas du point de vue développemental. Or les scores subjectifs obtenus sur ces textes, nous donnent aussi l'occasion d'observer le développement de la fluence, et particulièrement de la prosodie, par les scores des dimensions de la fluence du CE1 à la 5e. Les scores moyens à chaque dimension devraient augmenter à chaque niveau. Nous nous attendons cependant à un plafonnement rapide sur les scores de vitesse et décodage. A l'inverse, les scores d'expressivité devraient connaitre une croissance plus régulière.

Méthodologie - analyse statistique

Les données utilisées sont celles présentées au début de ce chapitre. Pour la suite de cette analyse, nous utiliserons pour chaque sujet le score moyen entre les deux textes et les trois évaluateurs experts. Les scores sont donc continus et compris entre 1 et 4 pour chaque dimension.

Toutes les analyses statistiques ont été réalisées avec le logiciel R (Team, 2019). L'effet du niveau a été testé en utilisant des modèles linéaires multiniveaux et en utilisant la classe comme paramètre aléatoire de niveau 2, avec la fonction lmer du package lme4 (Bates *et al.*, 2015). En effet les 295 enfants sont répartis dans 14 classes différentes. L'effet aléatoire classe est significatif pour toutes les dimensions de fluence (indice de corrélation interclasses > 0.05). Nous avons comparé par une ANOVA le modèle vide et le modèle avec le niveau comme effet fixe. Pour chaque analyse, nous rapportons le $\Delta\chi^2$(degré de liberté) et la p-value. Quand l'effet du niveau est significatif, une analyse post-hoc a été menée pour déterminer les différences entre les niveaux deux à deux. Nous avons utilisé le test HSD de Tukey, pour cette analyse nous rapportons la p-value la plus élevée parmi les tests deux à deux.

Résultats

La figure 3.3 présente l'évolution des scores avec le niveau scolaire pour chaque dimension.

Score total Le niveau a un effet significatif sur le score total à l'échelle EMDF ($\Delta\chi^2(5) = 72$, $p < .001$). Le score total augmente avec le niveau scolaire. Les comparaisons des scores par niveau montrent quatre étapes dans l'évolution. Les CE1 ont un score bien inférieur ($M = 8.42$, $SD = 2.05$) aux autres niveaux ($p < .001$). Le score augmente ensuite significativement en CE2 ($M = 10.75$, $SD = 2.85$) ($p < .001$), puis en CM1 ($p < .001$) avant de se stabiliser jusqu'en 6^{e} autour d'une moyenne de 12.72 ($SD = 1.71$). Il augmente ensuite légèrement de nouveau en 5^{e} ($p < 0.048$) pour atteindre 13.92 ($SD = 1.35$). Il est également intéressant de noter que la variabilité inter-individuelle pour ces scores totaux en CE1 et CE2 est assez importante, les écarts-types sont respectivement de 2.05 et 2.85, par rapport aux niveaux scolaires supérieurs, avec un écart-type autour de 1.5. Comme attendu, les scores de fluence augmentent avec le niveau scolaire et semblent plafonner à partir du CM1.

Score d'expressivité On observe un effet significatif du niveau scolaire sur le score d'expressivité ($\Delta\chi^2(5) = 38.7$, $p < .001$). Le score d'expressivité augmente progressivement avec le niveau scolaire. Les CE1 ont un score moyen ($M = 2.07$, $SD = 0.68$) inférieur aux autres niveaux ($p < .018$). Le score augmente en CE2 ($M = 2.46$, $SD = 0.85$) ($p < .002$). Il se stabilise ensuite à partir du CM1 avec un score moyen de 3.01 ($SD = 0.59$). Le score d'expressivité ne plafonne donc pas mais semble stagner autour de 3 assez rapidement. Le score maximal d'expressivité est donc difficile à atteindre pour tous les élèves. Il faut également noter que la variabilité inter individuelle évolue peu pour l'expressivité à partir du CM1 (entre 0.57 et 0.63). Elle est cependant plus élevée en CE2. Nous reviendrons sur cet aspect dans la discussion.

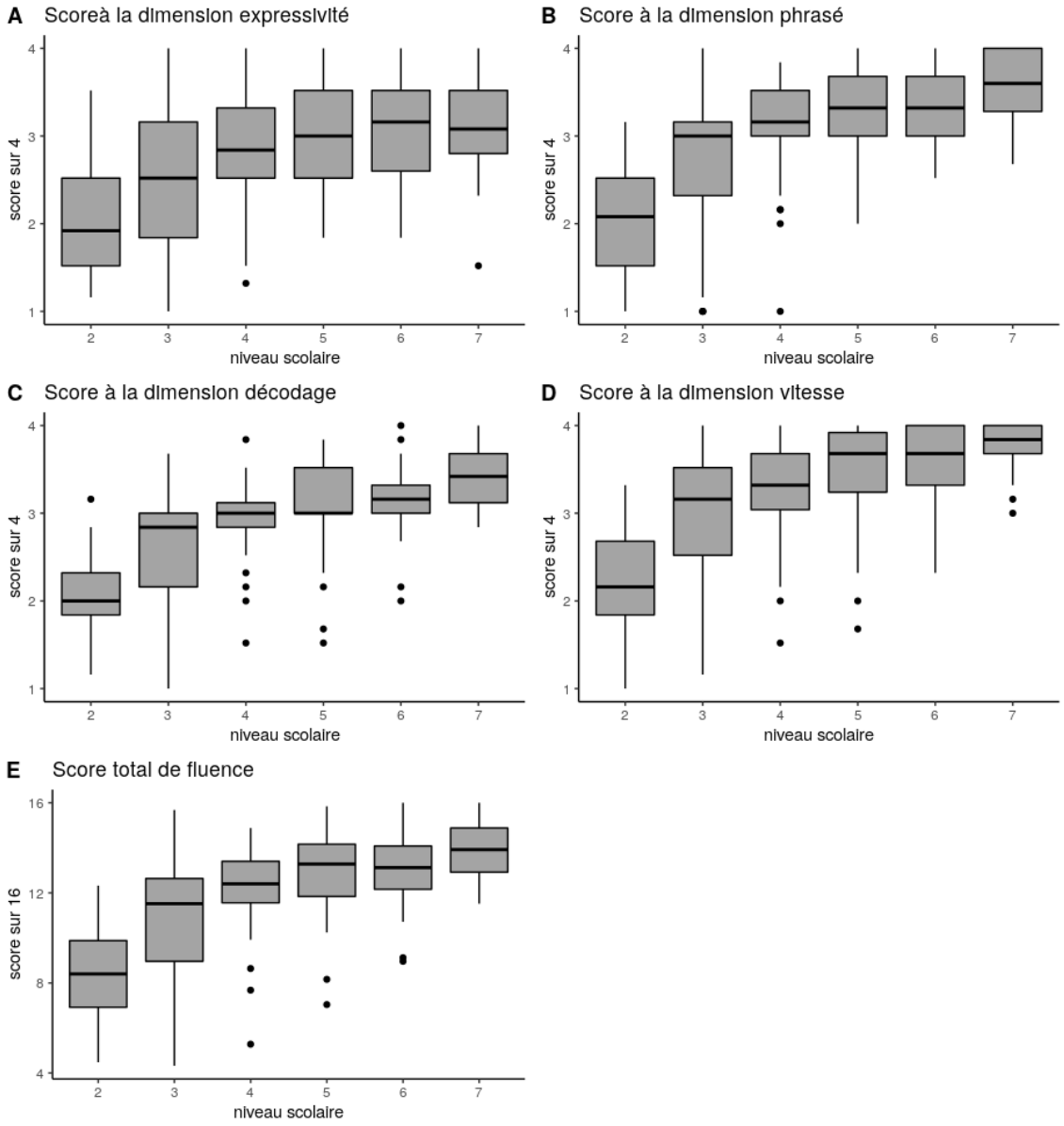


FIGURE 3.3 – Boxplot des scores aux différentes dimensions de l'EMDF en fonction du niveau scolaire

A. Expressivité B. Phrasé C. Décodage D. Vitesse. E. Score total

Score de phrasé On observe un effet significatif du niveau scolaire sur le score de phrasé (($\Delta\chi^2(5) = 74.9$, $p < .001$). Comme pour les autres dimensions et le score global, le score de phrasé augmente avec le niveau scolaire. On peut observer les mêmes différences entre niveaux que pour le score global. Les CE1 ont un score moyen faible (M = 2, SD = 0.63). Ce score augmente significativement en CE2 pour attendre 2.71 (SD = 0.80) ($p < .001$), puis de nouveau en CM1 pour arriver à 3.10 (SD = 0.51) ($p < .001$). Il reste ensuite stable autour de 3.24 (SD = 0.46) en CM2 et 6e, puis augmente de nouveau significativement en 5e pour atteindre 3.58 (SD = 0.42). On peut alors supposer que le score moyen atteint un plateau proche du score maximal sur ces textes. Une fois encore, il est intéressant de noter l'évolution de la variabilité inter-individuelle pour chaque niveau scolaire, avec une variabilité plus élevée en CE1 et surtout en CE2, que pour les autres niveaux. On peut également voir apparaitre sur le graphique B de la figure 3.3 des valeurs marginales en CM1 et CM2, avec une diminution de la variabilité.

Score de décodage On observe un effet significatif du niveau scolaire sur le score moyen à la dimension décodage ($\Delta\chi^2(5) = 69.4$, $p < .001$). Le score augmente avec le niveau scolaire. De 2.12 (SD = 0.41) en CE1, le score augmente significativement jusqu'à 2.65 (SD = 0.58) en CE2 ($p < .001$). L'augmentation est ensuite plus lente jusqu'en CM1 (M = 2.95, SD = 0.58) ($p = 0.012$). Le score moyen se stabilise ensuite à 3.10 (SD = 0.47) en CM2 et 6e. Une fois encore les 5es ont un score moyen significativement plus élevé que les autres niveaux scolaires à 3.42 (SD = 0.38) ($p < 0.006$). Le score moyen atteint en 5e semble aussi plafonner pour ces textes. Ceux-ci étant de niveau primaire, il est logique qu'ils ne présentent pas de difficulté de décodage et donc que les scores soient trés bons en fin de primaire et au collège. Pour la dimension de décodage également, on retrouve une variabilité inter-individuelle plus forte chez les CE2 (SD = 0.58) que chez les CE1 (SD = 0.41) et les lecteurs plus agés (SD < 0.50). Comme pour le phrasé, on peut observer sur la figure 3.3.C que, du CM1 à la 6e, la diminution de la variabilité inter-individuelle à la dimension décodage s'accompagne d'une augmention du nombre de valeurs marginales.

Score de vitesse On observe un effet significatif du niveau sur la vitesse moyenne ($\Delta\chi^2(5) = 77.3$, $p < .001$). Celui-ci augmente avec le niveau scolaire. Il est faible en CE1 (M = 2.25, SD = 0.58) puis augmente fortement en CE2 à 2.95 (SD = 0.82, $p < .001$). Le score moyen continue d'augmenter plus faiblement jusqu'en CM1 ($p < 0.04$), niveau à partir duquel il se stabilise jusqu'en 6e autour de 3.46 (SD = 0.50). Le score moyen atteint finalement 3.80 (SD = 0.27) en 5e. Comme pour le décodage, ce texte ne présentant pas de diffciulté particulière pour des élèves de fin de primaire et de collège, ces derniers atteignent une vitesse de lecture optimale, équivalente à de la parole spontanée, assez rapidement. On pourra noter, une fois encore, l'importante variabilité inter individuelle en CE2 (SD = 0.82) par rapport aux autres niveaux.

Discussion

Comme attendu, les scores moyens à l'échelle EMDF, ainsi qu'à chacune de ses dimensions, augmentent avec le niveau scolaire. On observe un effet plafond pour la dimension de vitesse qui atteint un score moyen trés proche du score maximal en 5e. Pour les scores de phrasé et de décodage, on observe également un effet plateau des scores moyens supérieurs à 3.5 à partir du CM1. Pour la dimension expressivité, le score moyen suit le même type d'évolution rapide en début d'apprentissage, pour arriver à un plateau à partir du CM1. Cependant, dans le cas de l'expressivité,

ce plateau se situe plutôt autour de 3, et n'atteint pas le score maximum. On constate donc que de toutes les dimensions de l'échelle de fluence, l'expressivité est la dimension acquise le plus tardivement et que cette acquisition est plus lente. En effet, le score moyen d'expressivité augmente lentement en début d'apprentissage, contrairement à ceux de décodage, vitesse et phrasé qui augmentent rapidement dès le CE1, pour s'approcher du score maximum dés le CM1. Considérant que les textes utilisés ici sont de niveau primaire, il est normal que des scores moyens supérieurs à 3 soient rapidement atteints. La dimension expressivité semble donc avoir un profil d'évolution différent des autres dimensions de la fluence. Ainsi des scores "experts" aux dimensions de décodage et de vitesse, mais également de phrasé, ne semblent pas suffisants pour atteindre un score "expert" en expressivité. Ces résultats vont dans le sens d'une acquisition du décodage et de l'automaticité, nécessaire, mais pas suffisante, à l'acquisition de la prosodie en lecture. Il serait donc intéressant de chercher les liens entre les différentes dimensions de l'EMDF et d'autre compétences en lecture. En effet dans la littérature, les études explorant le lien prosodie-compréhension utilisent uniquement le score global à l'échelle de fluence, sans s'attarder sur les particularités de chaque dimension. Or, comme nous venons de le voir, si le profil global d'évolution est relativement similaire pour chaque dimension, chacune présente ses particularités propres.

Le deuxième point important à soulever est la différence dans la variabilité inter-individuelle entre les niveaux. On constate que la variabilité inter-individuelle dans les scores moyens est beaucoup plus importante chez les plus jeunes. Cette constatation peut s'expliquer par les démarrages trés variables des enfants dans l'acquisition de la lecture. En effet en CE1 et CE2, le décodage et l'automaticité sont encore en cours d'acquisition pour certains enfants, alors que pour d'autres ces compétences sont acquises depuis le CP. Au delà du CE2, l'automaticité semble acquise par une grande majorité des enfants. A partir du CM1, la variabilité entre enfants diminue donc pour les compétences de décodage et de vitesse. On observe le même effet sur la dimension de phrasé alors qu'il est beaucoup moins présent sur l'expressivité. L'expressivité semble, là encore, présenter des caractéristiques différentes des autres dimensions. La variabilité est beaucoup plus régulière entre les niveaux, ce qui va dans le sens d'une dimension dont l'acquisition diffère des trois autres, liée à des compétences autres que le décodage et l'automaticité.

Finalement, la variabilité inter-individuelle est plus importante pour les CE2 que pour les CE1. Il est possible que les CE2, dont les compétences sont encore fragiles, présentent des stratégies plus variées : certains privilégient la vitesse au détriment de la précision, du phrasé et de l'expressivité, d'autres à l'inverse vont préférer lire plus lentement pour être précis et avoir un bon phrasé. Cette diversité de stratégies pourrait expliquer des scores trés variables dans toutes les dimensions. Une autre hypothése tient à un biais possible dans l'échantillon de CE1. En effet, une grande majorité des refus d'autorisation des parents pour la participation à l'étude provient de parents de CE1. D'aprés les enseignants, ces refus concernent des enfants non-lecteurs ou trés faibles lecteurs. Ainsi la cohorte de CE1 n'est pas aussi représentative des écoles que pour les autres niveaux. En effet, les sujets les plus faibles, non-lecteur en CE1, le sont à partir du CE2 même si leur niveau reste faible. Ce pourrait être cette frange de faibles lecteurs qui manquent à la cohorte de CE1, réduisant de fait la variabilité réelle qu'on peut observer dans les niveaux de lecture de CE1. Ces trés faibles lecteurs apparaissent d'ailleurs probablement sous la forme de valeurs marginales plus nombreuses en CM1 quand la variablité diminue, comme montré dans la figure 3.3.

Résultats majeurs **Étude 1**

L'EMDF, avec une méthodologie stricte, peut être un outil fiable et fidèle quels que soient les évaluateurs, pour peu qu'ils aient été suffisamment entrainés à l'utiliser.

Les dimensions phrasé et expressivité apporte de l'information supplémentaire par rapport au NMCLM

Pour chacune des dimensions de l'EDMF, le developpement est trés rapide en CE1 et CE2 puis ralentit pour plafonner à partir du CM2, à des niveaux qui ne sont pourtant pas encore ceux d'un lecteur expert.

L'expressivité se démarque comme la dimension obtenant les scores les plus faibles, et plafonne à des valeurs notablement inférieures au score attendu d'un lecteur expert.

3.2 Étude 2 : Étude du lien entre évaluation subjective de la prosodie et compréhension écrite

Dans la littérature, la corrélation entre prosodie et compréhension a été établie dans plusieurs langues en utilisant la MDFS. Ces études prennent toujours en compte le score total, score de fluence, et non pas les dimensions spécifiques à la prosodie, expressivité et phrasé. Comme nous l'avons montré dans l'étude précédente, l'expressivité et le phrasé se développent différemment. Il serait donc intéressant de voir quel est l'apport spécifique de chacune de ses dimensions dans les scores de compréhension écrite. Dans cette deuxième étude, nous utiliserons les données du corpus transversal pour, dans un premier temps, retrouver le lien entre fluence et compréhension. Dans un deuxième temps, nous irons plus loin en nous intéressant au lien entre la compréhension écrite et l'expressivité et le phrasé. Les dimensions prosodiques de la fluence peuvent-elles être prédites par la compréhension écrite, et les scores de compréhension écrites peuvent-ils être prédits par les dimensions prosodiques de la fluence ? Cette étude a été présentée lors d'une communication orale à la conférence annuelle de la SSSR (Society for Scientific Study of Reading) à Toronto le 24 juillet 2019.

3.2.1 Méthodologie

Données expérimentales

Fluence et prosodie Nous utilisons ici les évaluations subjectives des enregistrements des deux textes du corpus transversal, réalisées par trois évaluateurs experts en suivant le protocole présenté dans l'étude précédente. Le score retenu par enfant est le score moyen des deux textes et des trois évaluateurs, soit un score continu, compris entre 1 et 4 pour chacune des quatre dimensions de l'EMDF, et un score continu entre 4 et 16 pour le score de fluence, score total à l'EMDF.

Compréhension Nous utilisons ici deux types de compréhension : compréhension globale et compréhension en-ligne. Ces deux tests sont présentés dans le chapitre 2. Le score de compréhension globale est un taux de réussite aux questions. Il est compris entre 0 et 1. Le score de compréhension en-ligne est calculé à partir de la différence entre le NMCLM dans le texte signifiant et dans le texte non-signifiant. Pour ramener ce score entre 0 et 1, il est divisé par le NMCLM dans le texte signifiant. Les scores moyens, écart-type, minimum et maximum pour chaque niveau sont présentés dans le tableau 3.11.

Controles Finalement, afin d'évaluer le lien direct entre fluence et prosodie en lecture et compréhension, nous controlerons également les scores de vocabulaire, vitesse de lecture, raisonnement non-verbal et compréhension orale. Ces tests sont présentés dans le chapitre 2.

Analyses statistiques

Toutes les analyses statistiques ont été réalisées avec le logiciel R (Team, 2019). Pour déterminer le lien entre prosodie et compréhension nous avons utilisé un modèle général linéaire multiniveau avec la classe comme paramètre aléatoire de niveau 2. Les paramètres utilisés dans le modèle sont les suivants :

Table 3.11 – Scores moyens par niveau pour les tests de fluence, compréhensions et vocabulaire.

niveau	CE1	CE2	CM1	CM2	6e	5e
Fluence	.53(.13) .28-.77	.67(.18) .27-.98	.77(.11) .33-.93	.80(.11) .44-.99	.82(.10) .56-1	.87(.08) .72-1
Phrasé	.5(.16) .25-.79	.68(.20) .25-1	.78(.13) .25-.96	.82(.12) .50-1	.83(.11) .63-1	.89(.10) .67-1
Expressivité	.52(.17) .29-.88	.61(.21) .25-1	.73(.16) .33-1	.75(.14) .46-1	.75(.15) .46-1	.78(.14) .38-1
NMCLM	28(22) 28-124	80(31) 24-159	99(29) 22-168	118(42) 51-224	116(29) 53-174	143(38) 76-224
Vocabulaire	.49(.14) .23-.77	.55(.14) .22-.87	.63(.15) .33-.92	.72(.14) .42-.97	.66(.16) .25-.87	.72(.17) .35-1
CO	.50(.21) 0-.92	.65(.21) .18-1	.72(.20) .17-.92	.75(.18) .17-1	.61(.26) .08-1	.71(.22) .17-1
CE globale	.70(.23) .12-1	.71(.26) .11-1	.63(.20) .10-1	.65(.19) 0-1	.54(.23) .20-1	.70(.18) .30-1
CE en ligne	.15(.12) 0-.62	.17(.11) .02-.48	.20(.13) 0-.76	.19(.10) .03-.41	.17(.12) 0-.60	.15(.11) 0-.44

Note : CE : compréhension écrite, CO : compréhension orale, NMCLM : nombre de mots lus par minute dans le texte signifiant

- niveau
- Fluence : score globale à l'EMDF
- Expressivité : score d'expressivité à l'EMDF
- Phrasé : score de phrasé de l'EMDF
- Vitesse : NMCLM sur texte signifiant Evaléo
- Vocabulaire
- Compréhension orale
- Raisonnement non verbal
- Compréhension écrite globale
- Compréhension écrite en-ligne

Les interactions des paramètres les plus explicatifs avec le niveau et le NMCLM ont également été ajoutées dans les régressions. Cinq modèles ont été réalisés : deux pour expliquer les score de compréhension globale et en-ligne et trois modèles pour expliquer la fluence et ses dimensions de phrasé et d'expressivité. Les scores de fluence, dimensions de la fluence, vocabulaire et compréhensions, ont été ramenés entre 0 et 1. Pour les scores de fluence et ses dimensions, nous avons effectué une transformation de Smithson et Verkuilen (2006) afin de prendre en compte les possibles zéros qui pourraient poser problème dans le modèle. Nous avons ensuite réalisé les régressions bêta en utilisant la fonction glmmadmb du package *glmmADMB* (Fournier *et al.*, 2012). Nous avons réalisé des régressions multiples, en comparant par une anova les modèles successifs et en conservant les paramètres les plus significatifs à chaque étape, avant de calculer les coefficients de régression et les coefficients de détermination pour chaque variable significative du modèle. Nous rapportons la part de variance propre des variables d'intérêt, c'est à à dire leur part de variance, une fois retirée la variance expliquée par les variables contrôles.

TABLE 3.12 – Comparaison des modèles multiniveaux pour le score global de fluence et les scores aux dimensions de phrasé et d'expressivité

	fluence	**Score phrasé**	**expressivité**
Effets fixes			
Contante	-1.48(0.17)***	-1.86(0.25)***	-1.62(0.32)***
Compréhension orale			0.064(0.019)***
Vocabulaire	0.031(0.0065)***	0.049(0.009)***	
NMCLM	0.0096(0.0014)***	0.0097(.0002)***	0.011(0.002)***
Niveau 3	-0.14(0.23)	-0.13(0.33)	-0.26(0.44)
Niveau 4	1.14(0.29)***	1.33(0.40)***	1.67(0.55)**
Niveau 5	0.86(0.25)***	0.84(0.34)*	1.02(0.44)*
Niveau 6	0.54(0.34)	0.26(0.47)	1.12(0.61)
Niveau 7	0.56(0.39)	0.62(0.55)	0.75(0.36)
NMCLM*niv 3	0.0025(0.0018)	0.0031(0.0025)	0.0015(0.031)
NMCLM*niv 4	-0.0047(0.0018)*	-0.0053(0.0025)*	-0.0081(0.0033)*
NMCLM*niv 5	-0.0036(0.0016)*	-0.0028(0.0022)	-0.0065(0.0028)*
NMCLM*niv 6	-0.0014(0.0019)	0.0001(0.0027)	-0.0061(0.0034)
NMCLM*niv 7	-0.0013(0.0019)	-0.0003(0.0026)	-0.0055(0.0031)
Effets aléatoires			
niveau 2 (classe)	4.10-9(6.33.10-5)	1.17e-8(.00011)	0.012(0.11)
Qualité de l'ajustement			
AIC	-662	-578	-313
adjR2	.839	.783	.558

Note : * : p < .05 , ** : p < .01, *** : p < .001

3.2.2 Résultats

Fluence

Dans ce smodèles nous testons les effets de la compréhension écrite sur le score global de fluence, et les score de phrasé et d'expressivité. Les résultats des régressions multiples sont présentés dans le tableau 3.12. Le schéma de la figure 3.4 synthétise les modèles issus de ces régressions en présentant les effets fixes simples significatifs sur les scores de fluence, phrasé et expressivité.

Une fois controlé le vocabulaire et le NMCLM, la regression multiple sur le score total de fluence ne fait apparaitre aucun effet des scores de compréhension. Nous avons ensuite évalué l'interaction des variables explicatives avec le niveau. Cette interaction est significative uniquement pour le NMCLM ($p < .001$). Avec une rupture de pente entre CE2 et CM1, on observe l'importance du NMCLM pour la fluence dans les premières années d'apprentissage, influence diminuant dans le temps. On peut également noter que le modèle proposé ici est un modèle avec un bon coefficient de détermination ($R^2 = 0.84$). Aucune variable de compréhension, ni orale ni écrite, n'a d'effet significatif sur ce score global de fluence.

Pour aller plus loin, nous avons répété ces analyses en nous intéressant uniquement aux dimensions prosodiques de la fluence, soit au score de phrasé et au score d'expressivité. Les résultats des

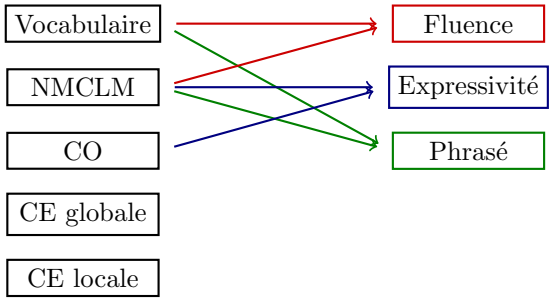


FIGURE 3.4 – Modéle des liens entre fluence, phrasé et expressivité avec la compréhension. Seuls les effets fixes simples significatifs sont représentés ici. Chaque couleur correspond à un modèle.

regressions multiples sont présentés dans le tableau 3.12. Le schéma de la figure 3.4 synthétise les modèles issus de ces regressions en présentant les effets fixes simples significatifs. Pour le phrasé, les variables ayant un effet fixe simple sont les mêmes que pour le score global de fluence : une fois controlé le vocabulaire et le NMCLM, les scores de compréhension n'ont aucun effet significatif. Les effets d'interaction avec le niveau sont significatifs pour le NMCLM. avec une pente élevée pour les premières années d'apprentissages, CE1 et CE2, qui diminue fortement en CM1. Il est intéressant de noter que le modèle proposé ici pour le phrasé, a un pouvoir explicatif un peu plus faible que pour le score global de fluence, avec un R^2 ajusté de 0.78. Ce modèle reste cependant un bon modèle pour nos données expérimentales.

Les résultats obtenus pour l'expressivité diffèrent en partie des résultats précédents. Une fois le NMCLM et la compréhension orale controlés, on n'observe aucun effet fixe simple de la compréhension écrite. On retrouve l'interaction entre NMCLM et niveau comme pour le phrasé. On peut donc noter que la compréhenion intervient dans notre modèle d'expressivité, mais uniquement sous la forme compréhension orale. La compréhension écrite n'a aucun effet significatif dans ce modèle. Le modèle proposé ici pour l'expressivité, a un pouvoir explicatif plus faible que pour le score de fluence et de phrasé, avec un R^2 ajusté de 0.56.

Compréhension

Dans ces modèles, nous testons les effets dde la fluence et de ses dimensions de phrasé et d'expressivité sur la compréhension écrite.

Les résultats des regressions multiples multi-niveaux pour la compréhension globale sont présentées dans le tableau 3.13. Le schéma de la figure 3.5 synthétise les modèles issus de ces regressions en présentant les effets fixes simples significatifs.

Une fois controlés le vocabulaire, le raisonnement non-verbal et la compréhension orale, la régression multiple sur le score de compréhension globale fait apparaitre l'effet fixe simple significatif du phrasé. Le tableau 3.13 présente les coefficients et la contribution des variables au pouvoir explicatif du modèle. On peut noter la contribution propre du score de phrasé et l'absence d'effet significatif du NMCLM. Nous avons ensuite évalué l'interaction des variables explicatives avec le niveau. Cette interaction est significative uniquement pour le vocabulaire ($p < .001$). On peut également noter que le modèle proposé ici est un modèle avec un coefficient de détermination assez

TABLE 3.13 – Comparaison des modèles multiniveaux pour le score au test de compréhension globale.

	Modèle 1	**Modèle 2**
Effets fixes		
Constante	-2.15(0.52)***	-2.33(0.52)***
Raisonnement NV	0.025(0.008)**	0.023(0.008)**
Compréhension orale	0.067(0.029)*	0.061(0.029)*
Vocabulaire	0.13(0.04)***	0.11(0.04)**
Niveau 3	-0.016(0.75)	-0.17(0.74)
Niveau 4	1.04(0.74)	0.69(0.75)
Niveau 5	0.46(0.81)	0.12(0.81)
Niveau 6	-0.69(0.9)	-1.06(0.91)
Niveau 7	-0.37(0.83)	-0.91(0.86)
Vocabulaire*niv 3	-0.01(0.05)	-0.015(0.047)
Vocabulaire*niv 4	-0.12(0.04)**	-0.11(0.04)*
Vocabulaire*niv 5	-0.099(0.044)*	-0.091(0.044)*
Vocabulaire*niv 6	-0.039(0.05)	-0.033(0.049)
Vocabulaire*niv 7	0.048(0.045)	-0.035(0.045)
Phrasé		1.07(0.46)*
Effets aléatoires		
Niveau 2 (classe)	8.2e-9(9e-5)	2.1e-9(4.5e-5)
Qualité de l'ajustement		
AIC	-176	-179
adjR^2	.348	.359
ΔadjR^2	-	.011

Note : * : p < .05 , ** : p < .01, *** : p < .001

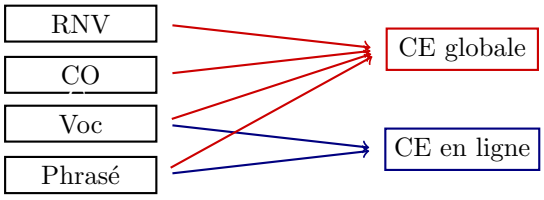


FIGURE 3.5 – Modéle des liens entre compréhension et fluence. Seuls les effets fixes simples significatifs sont représentés ici. Chaque couleur correspond à un modèle

TABLE 3.14 – Comparaison des modèles multiniveaux pour le score au test de compréhension en-ligne

	Modèle 1	**Modèle 2**
Effets fixes		
Constante	-2.19(0.18)***	-2.75(0.21)***
Vocabulaire	0.075(0.011)***	0.059(0.011)***
NMCLM	-0.004(0.001)***	-0.0069(0.0008)***
Phrasé		1.93(0.37)***
Effets aléatoires		
Niveau 2 (classe)	0.0023(0.049)	1.4e-8(1.1e-4)
Qualité de l'ajustement		
AIC	-504	-529
adjR^2	.233	.299
ΔadjR^2	-	.063

Note : * : p $<$ .05 , ** : p $<$ .01, *** : p $<$ .001

faible ($R^2 = 0.359$). Les variables choisies ici sont donc insuffisantes pour expliquer les scores de compréhension.

Pour la compréhension en-ligne, les résultats des régressions multiples multi-niveaux sont présentés dans le tableau 3.14. Le schéma de la figure 3.5 synthétise les modèles issus de ces regressions en présentant les effets fixes simples significatifs.

Une fois le vocabulaire controlé, la régression multiple sur le score de compréhension globale fait apparaitre l'effet simple significatif du phrasé. Le tableau 3.13 présente les coefficients et la contribution de chaque variable au pouvoir explicatif du modèle. On peut noter, comme pour la compréhension globale, la contribution propre du score de phrasé. Nous avons ensuite évalué l'interaction des variables explicatives avec le niveau et le NMCLM. Il n'y a pas d'interaction significative avec le niveau pour la compréhension en-ligne. Il y cependant une interaction significative entre vocabulaire et NMCLM (p $<$ .001). On peut également noter que le modèle proposé ici est un modèle avec un coefficient de détermination assez faible ($R^2 = 0.30$). Les variables choisies sont donc insuffisantes pour expliquer les scores de compréhension en-ligne.

3.2.3 Discussion

Dans cette étude, nous avons pu mettre en évidence le lien entre score de compréhension écrite et score de fluence mesuré par l'EMDF, et plus spécifiquement ses dimensions prosodiques. En effet, le score de phrasé explique une part de variance propre des scores de compréhension globale et en-ligne, au delà du vocabulaire, du raisonnement non-verbal et de la compréhension orale. Dans les scores à l'EMDF, seule la compréhension orale prédit une part de variance du score d'expressivité. Nous avons ainsi pu mettre en évidence, en français, le lien entre prosodie en lecture et compréhension présent dans plusieurs langues (cf chapitre 1). Au delà de ce lien, plusieurs points sont à noter dans cette étude.

Tout d'abord, il est interessant de s'attarder sur le rôle du NMCLM. En effet, le NMCLM est le principal prédicteur du score de fluence mesuré avec l'EMDF. On retrouve ici la corrélation entre NMCLM et score à l'EMDF, mentionnée dans l'étude 1. Pour rappel, cette corrélation diminue avec le niveau scolaire. On retrouve le même phénomène dans notre modèle avec l'interaction entre NMCLM et niveau scolaire. Il est également intéressant de noter que l'effet du NMCLM et l'interaction avec le niveau sont également présents sur les deux dimensions prosodiques. Cela vient renforcer l'idée de la necessité de l'acquisition de l'automaticité pour le développement de la prosodie. Cependant le NMCLM n'apparait pas comme prédicteur des scores de compréhension, contrairement au score subjectif de fluence, via le phrasé. Ces deux points viennent étayer la necessité d'élargir le type de mesures effectuées pour évaluer les compétences en lecture. Il est nécessaire d'aller au-delà de la mesure du NMCLM. L'évaluation subjective via une échelle de fluence incluant des dimensions prosodiques apparait ainsi être un bon complément pour avoir une vue plus générale des compétences en lecture.

Le score de compréhension en-ligne traduit la façon dont le sujet s'appuie sur le sens du texte pour anticiper le mot ou la phrase suivante et ainsi lire plus rapidement le texte. On retrouve le phrasé comme prédicteur de ce score, comme pour la compréhension écrite globale. Cet impact du phrasé pourrait s'expliquer par l'analyse syntaxique nécessaire à la fois au phrasé et à la compréhension en-ligne. En effet, comme vu précédemment, un phrasé correct repose sur une reconnaissance de la structure syntaxique du texte. Ainsi les enfants qui sont capables d'extraire la structure syntaxique du texte pour la mettre en valeur par le phrasé, peuvent également s'en aider pour anticiper le sens du texte à venir.

Finalement, l'évaluation de la compréhension écrite est une question complexe, particulièrement pour la comparaison entre différents niveaux en employant, comme ici, des textes différents. Nous n'avons en français aucune batterie de tests standardisés de compréhension écrite, conçue pour l'évaluation en milieu scolaire, comme il peut en exister en anglais par exemple (Marston, 1989, par exemple). Nous avons choisi d'évaluer la compréhension globale avec un texte en lecture silencieuse, suivi de questions. Ce type d'évaluation implique d'autre compétences que la compréhension, telles que la mémoire de travail ou l'attention, qui peuvent fortement moduler les résultats notamment chez les plus jeunes. Ainsi il pourrait être intéressant dans une prochaine étude de compléter ces tests de compréhension par d'autres types de test, comme des questions orales sur le texte qui vient d'être lu à voix haute, demander à l'enfant de raconter l'histoire qu'il vient de lire, ou encore des Maze tests qui mesurent la compréhension écrite au niveau plus local de la phrase ou du paragraphe.

Une autre limite des modèles présentés dans cette étude est l'effet plafond qu'on peut observer dans l'évaluation de la fluence et des dimensions prosodiques. En effet, les textes utilisés pour évaluer la prosodie ici sont de niveau primaire et les scores plafonnent dés le CM2 sur certaines dimensions. Il est donc possible que nous ne captions pas toutes les variables pouvant expliquer la fluence et la compréhension dans les niveaux scolaires les plus hauts. Il serait alors intéressant de renouveller cette étude au collège en utilisant pour la lecture orale des textes plus adaptés au niveau scolaire des adolescents. Ce nouveau protocole serait susceptible de mieux capter les variations dans les liens entre prosodie et compréhension au collège.

Résultats majeurs **Étude 2**

La prosodie, à travers le score à la dimension phrasé, apporte une part de variance propre des scores de compréhension écrite globale et en-ligne.

Le NMCLM n'apporte aucune information sur la compréhension écrite.

La compréhension écrite ne prédit pas les scores de fluence ou ses dimensions.

Le niveau et le NMCLM modulent l'impact des prédicteurs.

3.3 Conclusion

Dans ce premier objectif, nous avons posé la question de l'utilisation d'une évaluation subjective pour caractériser le développement de la prosodie en lecture et son lien avec la compréhension écrite. L'échelle subjective EMDF est une échelle traduite et adaptée de l'anglais. Avec les textes utilisés pour la cohorte testée, l'EMDF est un outil sensible, fidèle et valide jusqu'en fin de primaire. Même si les textes utilisés dans ces études entrainent un effet plafond pour les collégiens, ils permettent quand même d'identifier des lecteurs en difficulté en 6^{e} et 5^{e}. De plus, l'EMDF est utilisable quelles que soient les connaissances des juges en matière de lecture d'enfant. Elle nécessite cependant un entrainement suffisant à son utilisation, particulièrement dans le cas d'évaluateurs multiples.

L'étude 1 montre que les scores à toutes les dimensions de l'EMDF augmentent avec le niveau scolaire, trés fortement du CE1 au CE2, puis plus lentement jusqu'en CM1 et CM2. Dans le cadre des textes utilisés ici, les scores plafonnent à partir du CM2. L'évaluation subjective transversale n'est donc pas l'outil le plus adapté à l'étude du développement sur une période de temps longue. Elle est cependant suffisamment sensible pour les premières années pendant lesquelles l'évolution est rapide et permet de capter l'étape importante du CM1. Concernant le lien prosodie-compréhension, l'utilisation de l'EMDF dans l'étude 2 a permis de retrouver le lien entre fluence, prosodie et compréhension, présent dans la littérature dans d'autres langues que le français. En effet, nos modèles ont montré que le phrasé explique une part de variance propre des scores de compréhension écrite et que le score d'expressivité est lié à la compréhension orale. L'utilisation de l'évaluation subjective dans la modélisation de ce lien donne donc des résultats intéressants. Cependant, une fois encore, les interactions avec le niveau mettent en évidence l'effet plafond au collège et ainsi les limites d'utilisation d'un même texte pour des niveaux aussi différents que des CE1 et des 6es-5es.

Finalement, il est important de noter dans ces deux études le profil particulier de la dimension expressivité parmi les dimensions de la fluence. En effet, comme pour les autres dimensions, les scores plafonnent également à partir du CM2 mais sans atteindre un score maximal. La variabilité inter-individuelle est plus régulière dans le temps que pour les autres dimensions de la fluence. D'un point de vue du lien prosodie-compréhension, l'expressivité se démarque également du score total et du phrasé, étant la seule liée à la compréhension orale. L'expressivité se développe donc différemment des autres dimensions et met probablement en jeu des compétences plus larges que les simples décodage et automaticité. Ainsi il serait intéressant de poursuivre ces analyses dans une étude longitudinale, pour voir comment varient les relations entre ces différentes dimensions de la fluence et les compétences en lecture. Comme nous avons pu le voir ici, l'échelle subjective va présenter des limites dans ce type d'étude : variations dans les jurys, captation de variations fines dans les dimensions mesurées... Il pourra donc être nécessaire d'adapter notre outil d'évaluation de la prosodie en lecture à ce type d'étude.

Chapitre 4

Caractérisation du développement de la prosodie par l'analyse des pauses et de la respiration

Dans ce chapitre, nous allons aborder le développement de la prosodie en lecture d'un point de vue des mesures acoustiques. Comme mentionné dans le chapitre 1, les paramètres acoustiques traditionnellement exploités pour évaluer la prosodie en lecture sont les paramètres liés aux pauses (nombre, durée, grammaticalité) et les paramètres liés aux variations de F0. Nous nous intéressons ici à l'évolution de ces paramètres dans des lectures d'enfants du CE1 à la 5^{e} et à leur comparaison avec les spécificités des lectures adultes. Les paramètres de pauses sont peu variables d'un lecteur expert à un autre car fortement liés à la syntaxe et à la ponctuation (cf. chapitre 1). Il existe donc un objectif relativement normatif à atteindre pour ces paramètres. A l'inverse, les paramètres liés aux variations de F0 présentent des caractéristiques très variables d'un lecteur expert à un autre. L'évaluation de la mélodie par comparaison à des lecteurs experts est donc beaucoup plus complexe. Ainsi, nous avons choisi de nous concentrer sur les pauses dans ce chapitre. Nous reviendrons sur les paramètres associés aux variations de F0 dans le chapitre suivant.

Comme mentionné dans les chapitres précédents, les pauses dans la lecture d'un texte sont nécessaires à sa bonne compréhension. Elles permettent de souligner la structure syntaxique de l'énoncé. Elles peuvent aussi révéler une activité cognitive de décodage, par exemple à travers les pauses marquant des hésitations. Mais les pauses servent en premier lieu à respirer pendant la lecture. La bonne utilisation des pauses respiratoires nécessite une planification en accord avec la syntaxe du texte et une coordination respiration-syntaxe qui demande une anticipation de ces deux aspects. L'absence ou une mauvaise planification des pauses syntaxiques ou respiratoires, en créant des ruptures dans la syntaxe et le rythme du texte a un impact sur le jugement subjectif de l'auditeur et nuit à la bonne compréhension du texte. Dans ce chapitre, nous présentons dans une première étude le développement de cette planification des pauses du CE1 à la 5^{e}. Dans un deuxième temps, nous explorons les liens entre les paramètres de pauses et les jugements subjectifs réalisés dans le chapitre précédent. Finalement, nous nous intéressons au lien entre ces paramètres de pauses et la compréhension écrite.

4.1 Étude 3 : Pauses et respiration du CE1 à l'adulte expert

Cette première étude utilise des outils d'analyse acoustique pour évaluer la prosodie des lectures d'enfants effectuées à voix haute. Elle se centre sur le développement de la planification des pauses et de la respiration en lecture. Elle a fait l'objet d'un article soumis au journal Reading and Writing : an Interdisciplinary Journal en juillet 2020, reproduit ci-dessous.

Godde, E., Bailly, G. & Bosse, M. L. (2020). Pausing and Breathing while Reading Aloud : Development from 2nd to 7th grade (soumis)

Pour faciliter le traitement en ligne des textes pour soi et pour autrui, le lecteur doit placer ses pauses, notamment respiratoires en accord avec la syntaxe du texte afin de ne pas perturber la compréhension. L'attention portée aux pauses permet de mieux comprendre le discours, mais également d'observer la gestion de l'activité cognitive et physiologique du lecteur. Chez les adultes, on observe une distribution multimodale de la longueur des pauses en lecture à voix haute (e.g. pauses longues et courtes) liée à une variété de fonctions : physiologique, syntaxique, hésitation. Les motifs de pauses sont très liés à la ponctuation et aux frontières syntaxiques majeures, ce qui demande une planification précise. L'acquisition de la planification des pauses chez les jeunes lecteurs a été peu étudiée. En espagnol et en anglais, le nombre et la durée des pauses ainsi que le nombre de pauses non grammaticales diminuent fortement durant les premières années d'apprentissage de la lecture, avec l'acquisition de l'automaticité. Cette compétence évolue ensuite avec l'acquisition des autres compétences en lecture. Les pauses respiratoires représentent 70% des pauses longues chez les adultes. Leur durée diminue quand la vitesse de lecture augmente, la planification est alors de plus en plus difficile. La coordination voix-respiration en parole se développe jusqu'à 10 ans. Le développement de la planification des pauses en lecture chez les apprentis lecteurs n'a, à notre connaissance, jamais été étudié bien que celle-ci soit révélatrice de la maturation de l'automatisation de la lecture. L'étude présentée dans cet article propose de décrire le développement de la planification des pauses et de sa coordination avec la respiration chez les enfants du CE1 à la 5^{e} en les comparant à des références adultes.

Dans cette étude, nous avons utilisé les enregistrements du texte 1 du corpus transversal. Ces enregistrements ont été alignés et les pauses annotées en fonction de leur caractère respiratoire ou non, grammatical ou non. Pour les pauses respiratoires, le délai inspiration-phonation (DIP) a également été mesuré. Sur 16895 pauses, 8900 sont des pauses respiratoires accompagnées d'une inspiration. Son début est identifié sur les traces capturées par les ceintures respiratoires et l'enregistrement audio des bruits de respirations. Des comparaisons de modèles linéaires multiniveaux, suivies d'analyses post-hoc de Tukey ont permis de déterminer l'effet du niveau scolaire et les différences significatives entre les niveaux pour chacun des paramètres étudiés : nombre et durée des pauses respiratoires et non respiratoires, nombre de pauses non grammaticales, nombre de pauses liées à la ponctuation, délai inhalation-phonation (DIP).

Nous avons ainsi pu mettre en évidence une bimodalité de la distribution de la durée des pauses pour chaque niveau scolaire. Les pauses ont donc été réparties en pauses courtes et pauses longues selon un seuil spécifique à chaque groupe. Les résultats des analyses ont permis de dégager quatre étapes dans le développement de l'acquisition des pauses. En CE1, les enfants font de nombreuses pauses, souvent non grammaticales. La planification est faible, et s'appuie en partie sur la ponctuation. A ce niveau là, la variabilité intra- et inter-sujets est grande. En CE2, le nombre et

la durée des pauses diminuent. La respiration commence à être planifiée. Cependant, on observe une augmentation du nombre de pauses non grammaticales, malgré l'appui sur la ponctuation. A partir du CM1, la maturation du contrôle du système respiratoire permet une meilleure planification. Le nombre de pauses diminuent fortement, particulièrement les pauses non grammaticales. L'appui sur la ponctuation est plus fort. Jusqu'en 6^{e}, les caractéristiques sont stables mais pas encore similaires à celles des lecteurs adultes. Finalement en 5^{e}, on peut retrouver les caractéristiques des lectures adultes, exceptée la durée moyenne des pauses qui demeure inférieure à celle des adultes.

Parmi les résultats remarquables, il est intéressant de mentionner la durée moyenne des pauses longues. Chez les CE1, celle-ci est équivalente à celle des adultes et supérieure à celle des autres enfants. Les raisons de cette variation sont cependant différentes pour les adultes et les CE1. Pour les CE1, les nombreuses pauses correspondent à une activité cognitive de décodage. La respiration étant moins planifiée, les DIP sont alors plus longs. Les adultes, à l'inverse, planifient ces pauses longues pour souligner la syntaxe du texte. Un autre fait intéressant est l'augmentation du nombre de pauses non grammaticales en CE2. Celle-ci pourrait être dûe à une accélération de la vitesse de lecture avec l'acquisition de l'automaticité, au détriment de la planification des pauses.
La première limitation de cette étude est l'utilisation d'une première lecture pour les analyses. Les enfants n'ont pas eu l'occasion de découvrir le texte avant les enregistrements. Il serait intéressant d'étudier également la planification des pauses après une première découverte du texte. De plus, nous n'avons pas distingué les pauses correspondant aux retours à la ligne dans la mise en page proposée. Une étude de ces pauses particulières, assortie éventuellement d'une étude en oculométrie, nous renseignerait sur l'anticipation permettant la planification ou l'inhibition des pauses. Finalement, les données transversales présentées ici nous permettent une description de la planification des pauses pour les différents niveaux. L'analyse des données longitudinales nous permettra de mieux comprendre l'évolution et les liens entre les différents paramètres.

Pausing and Breathing while Reading Aloud : Development from 2nd to 7th grade

Abstract

Pauses when reading aloud play an essential role in reading and listening comprehension. Among the various types of pauses, breathing pauses during oral reading are particularly essential. Their placement, frequency and duration tell us about breath and voice coordination as well as articulatory planning. These skills that develop over time, from the early stages of reading acquisition to skilled reading levels, are under-researched. The present study aims to describe how children develop the ability to use pauses during oral reading by comparing younger French speaking learners ranging from second to seventh graders to adults. Voice and breathing patterns of 295 children and 20 adults were recorded during oral reading. Specific focus was given to the frequency, duration and placement of breathing pauses and inhalation to phonation delay. Our results revealed that the youngest students took more ungrammatical pauses, breathed more frequently and produced longer inhalation to phonation delays. Children older than grade 4 were visibly more proficient at planning their pauses, particularly their breathing pauses, and tended to rely more on punctuation. Finally, pause planning observed for the grade 7 students was almost at the adult level. This developmental study suggests that children's acquisition of pausing patterns during oral reading can effectively be described in four steps that occur between the 2nd and 7th grades, and highlights the important role punctuation plays in planning breathing pauses and syntactic-prosodic breaks.

Keywords : reading prosody, breathing-voice coordination, reading development, pausing in reading

Introduction

Reading aloud involves numerous complex cognitive processes : letter recognition, grapheme-phoneme mapping, semantic retrieval, syntactic extraction, articulatory planning, etc. An important part of children's reading acquisition process involves learning to coordinate of all these processes. In the first years of reading acquisition, teachers focus primarily on skills involving automation of word decoding. Once a young learner's decoding skills become more automatic, their cognitive resources can be freed for other purposes such as developing comprehension. The ability to add prosodic breaks during oral reading of a given text requires engaging further cognitive resources. To read with adequate prosody, the reader must add appropriate intonation and plan pauses to effectively match the syntax and the meaning of the text (Kuhn *et al.*, 2010). In speech, pauses play an essential role in defining the communication purpose. They not only highlight the syntax of the discourse but allow the listener to understand meaning. The specific nature of pauses may also reveal the cognitive load of the speaker as, for example, longer pauses are typically associated with thinking and planning the next sequence (Brennan et Williams, 1995). Moreover, placement of pauses, their duration and frequency have been shown to be closely linked to text comprehension (Molomer *et al.* 2015; Arcand *et al.* 2014, in French). Lalain *et al.* (2016) proposed that, although silent, pauses can be an integral part of the message being conveyed and merit close attention. The present study aims to fill the gap of knowledge about the development of reading pauses in French children.

Pausing in adults In adults, pauses have been studied in different oral modalities, for example in reading and in spontaneous speech that lead to different pause configuration (Goldman *et al.*, 2010). Several authors showed a multimodal distribution of pause duration in several languages. Campione et Véronis (2002) showed a bimodality of pause duration in text reading, brief pauses (< 200ms) and medium pauses (200-1000ms), in five languages (French, German, English, Italian and Spanish). Pauses longer than 1000ms were only present in spontaneous speech. Goldman *et al.* (2010) also found a bimodal distribution of pauses in text reading, with short pauses around 100ms and longer ones measuring roughly 500ms. Bailly et Gouvernayre (2012) found that a French expert reader reading an audio book made short pauses (< 200ms) and three kinds of long pauses classified according to their location : sentence internal pauses (occuring at major syntactic boundaries or punctuation), end of sentence pauses and end of paragraph pauses. Lalain *et al.* (2014) (see also Grosjean et Deschamps 1975) classified pauses into three main categories according to their uses : syntactic pauses highlighting text comprehension, hesitation pauses revealing a high cognitive load, and physiological pauses that are mainly respiratory. Grosman *et al.* (2018) also linked the characteristics of pause patterns (i.e., duration and placement) to their discursive functions. This mapping depends mostly on the type of speech. In reading aloud, they observed two modes : short pauses linked to articulation processes and long pauses linked to cognitive processes : hesitation, syntactic and physiological. These authors focused their analysis on each category of long pauses considered separately rather than mean measures or overall pause occurrence. However, adults make very few long hesitation pauses. In adult reading, 60% of pauses occur at the end of sentences (Grosjean et Deschamps, 1975; Grosman *et al.*, 2018). These are long syntactic pauses often used by the reader to breathe. Pauses have also been shown to be longer when they occur between particularly long sentences or longer phrases than for shorter sentences (Zvonik et Cummins, 2003). Pausing patterns in adult reading has also been closely linked to the text punctuation (Martin, 2011; Campione *et al.*, 2002).

Pausing in children reading With mostly syntactic pausing at major boundaries and punctuation, adult pausing patterns rely on precise planning and online processing of the syntactical structure. Prosodic planning is considered a reliable marker of reading fluency and overall reading skills (Cowie *et al.*, 2002; Benjamin *et al.*, 2013) and is clearly linked to written comprehension (e.g. Álvarez-Cañizo *et al.*, 2015; Arcand *et al.*, 2014; Schwanenflugel *et al.*, 2004; Miller et Schwanenflugel, 2006). Children, on the other hand, appear to use pauses to aid their comprehension when reading difficult texts (Benjamin et Schwanenflugel, 2010) and may serve as an additional tool to enhance their online comprehension abilities. However, pauses planning is rarely explicitly taught to children when they are learning to read. To date, very few studies have investigated the development in younger learners of adult level pausing in reading (Godde *et al.* 2020a, for a review). Studies in English (Schwanenflugel *et al.*, 2004; Miller et Schwanenflugel, 2006) and Spanish (Álvarez-Cañizo *et al.*, 2018) showed similar learning development patterns in the two languages that began with decoding acquisition and gradually moved toward the acquisition of the other reading skills, such as reading fluency and comprehension. Pausal intrusions, also called ungrammatical pauses, are intra-sentential pause patterns occurring at syntactic boundaries, for example, between a noun and its specifier or adjective, after a preposition or inside a word. This type of pause is very frequently observed in young readers and tends to decrease rapidly from grade 1 to grade 3. Third graders, however, tend to produce more intra-sentential pauses and within-word pauses than adults ((Álvarez-Cañizo *et al.*, 2018). Pause frequency and their duration, both inter and intra-sentential, also decrease rapidly from grade 1 to grade 3, but still do not reach adult level at the end of the primary grade. Moreover, pausing patterns in oral reading depend on the individual's

reading skill level : compared to fluent readers in the same grade, non-fluent readers tend to make more frequent and longer pauses, and produce more ungrammatical pauses (Godde *et al.*, 2020b). As Grosjean et Collins (1979) reported in their study of pausing in adults, Breznitz (1990) showed that pausing patterns are also at least partially dependent on reading speed. Increased reading speeds in children corresponded to fewer and shorter pauses. Additionally, young children appear to rely more heavily on punctuation to plan their pauses than older, more experienced readers. Miller et Schwanenflugel (2008) proposed that this may be due to a decrease of motivation to use punctuation as reading skills improve. To our knowledge, no studies have investigated development of pauses in young readers older than primary school age but who have clearly not fully matured as skilled oral readers.

Pausing and breathing in reading Among the various types of pauses, breathing pauses are essential and incompressible. In both spontaneous speech and reading, in order to match the syntax of the text, the reader's breathing must be carefully planned to effectively balance both physiological anexample, between a noun and its specifier or adjective, after a preposition or inside a word. This type of pause is very frequently observed in young readers and tends to decrease rapidly from grade 1 to grade 3. Third graders, however, tend to produce more intra-sentential pauses and within-word pauses than adults (Álvarez-Cañizo *et al.*, 2018). Pause frequency and their duration, both inter and intra-sentential, also decrease rapidly from grade 1 to grade 3, but still do not reach adult level at the end of the primary grade. Moreover, pausing patterns in oral reading depend on the individual's reading skill level : compared to fluent readers in the same grade, non-fluent readers tend to make more frequent and longer pauses, and produce more ungrammatical pauses (Godde *et al.*, 2020b). As Grosjean et Collins (1979) reported in their study of pausing in adults, Breznitz (1990) showed that pausing patterns are also at least partially dependent on reading speed. Increased reading speeds in children corresponded to fewer and shorter pauses. Additionally, young children appear to rely more heavily on punctuation to plan their pauses than older, more experienced readers. Miller et Schwanenflugel (2008) proposed that this may be due to a decrease of motivation to use punctuation as reading skills improve. To our knowledge, no studies have investigated development of pauses in young readers older than primary school age but who have clearly not fully matured as skilled oral readers. d syntactic requirements.

Wlodarczak et Heldner (2017) showed that the syntactic structure is predominant over the physiological needs during spontaneous speech. Adults are able to continue talking as long as they have enough air in their lungs to complete the next utterance (Wang *et al.*, 2010). There seems to be no adaptation of the syntactic structure to the breathing needs. Breathing pauses, however, can be longer and less grammatical when the speaker hesitates. The cognitive load then increases and the syntax is less structured : consequently, anticipation on the next utterance becomes more difficult. Hesitation pauses are more often present during spontaneous speech when the discourse is less constrained than during oral reading and are almost non-existent in adult level reading, with less than 2% of inappropriate breathing pauses (Wang *et al.*, 2010).

In oral reading, the structure is constrained by the text. Bailly et Gouvernayre (2012) analyzed the pausing and breathing coordination of one French reader of audiobooks and observed that 70% of medium and long pauses were breathing pauses. They also showed that breathing noises are used as supplementary cues of the text structure, supporting the importance of planning breathing pauses in reading. Grosjean et Collins (1979) showed that high reading speed has an impact on breathing while reading. As reading speed increases, the frequency and duration of non-respiratory pauses initially decrease. In a second phase, if the reading rate increases, the number

and duration of the breathing pauses decrease further. Breathing pauses can be divided in three phases : pre-inspiration, inhalation and post-inspiration. When reading speed increases, the pre-inspiration phase tends to disappear. In adults, inhalation always occurs at the end of the breathing pause. According to Grosjean et Collins (1979), in reading measured at a standard rate, breathing patterns will be more dependent on preplanned pause patterns. But as the reading rate increases, physiological needs take over planning and patterns of breathing pauses are no longer strongly linked to the syntax but almost exclusively depend on inhalation demands.

Coordinating breathing and text reading thus presents a specific challenge in the context of developing pause planning skills in reading. Children must learn to balance between the need to breath and respecting the syntactic structure of the text. They need to learn to anticipate what's coming next in the text, and learn to use punctuation as visual cues to help place their inspiration during syntactic pauses. Breathing coordination during oral reading is particularly challenging for children to master and breathing-speech coordination has been shown to progressively develop between 3 and 10 years of age. As the number of syllables pronounced in one exhalation gradually increases, the level of difficulty coordinating breath inspiration with linguistic or communicative purpose decreases (Boliek *et al.*, 2009). It has also been shown that a boost in progress tends to occur between the ages 7 and 8 years (Hoit et al., 1990) and important inter individual variability in developing these skills is also at play (Boliek *et al.*, 2009). In short, for children beginning to read around 6 or 7 years of age, lung capacity and speech-breathing coordination are not yet fully matured.

Despite the critical role of pausing and breathing during reading, surprisingly little attention has been dedicated to describing the development of controlled breathing patterns in reading. Among the few studies that have investigated this aspect of oral reading, one study focused on children with and without asthma (Wiechern *et al.*, 2018) while another study children with and without dyslexia (Lalain *et al.*, 2014). For control children, these selected studies were only able to show that fewer but longer breathing pauses were produced during the reading of difficult text. The primary purpose of our study is to precisely characterize the development of pausing in reading from grade 2 to grade 7, by clearly differentiating breathing and non-breathing pauses. To do this, we focused on four core perspectives of pausing during reading. We began by considering pause duration and its eventual multimodality, as observed in adults and anticipated that this would reveal two separate modalities, short and long pauses, that develop in parallel with reading acquisition. Secondly, we focused on the number or frequency of pauses. According to literature, we hypothesized that the number of long pauses would decrease rapidly in the initial years of learning (reading acquisition) and that the number of ungrammatical pausal would be adult-like for pupils at who reached primary school age. Thirdly, we concentrated specifically on breathing pauses spotted by the use of two respiratory belts : their frequency, duration and their planning through the inhalation to phonation delay. Our hypothesis here was that the number of breathing pauses would decrease up to the end of primary grades when breathing patterns mature. Ungrammatical breathing pauses should be frequent in the first years but should quickly disappear. Concerning the duration of breathing pauses, we anticipated mostly longer pauses, and that the inhalation to phonation delay would decrease as reading skills developed. Finally, we considered the potentially important role punctuation marks play in terms of planning pauses during oral reading. Here we expected that younger children would rely more heavily on punctuation as signals to pause and breath than older children and adults, particularly when navigating longer, more complex sentences.

Method

This study took place in two primary schools and one middle school located in the urban area of Grenoble (France). Our protocol was approved by the ethics committee (CERGA, ethic committee of Université Grenoble Alpes) and the local representative of the French Ministry of Education.

Participants Audio recordings of oral readings were made for 323 students ranging from grade 2 to grade 7 from 14 classes in three different schools. Authorization release forms signed for by each participating student's parent(s), teacher and school director were obtained were obtained prior to the recordings. For the purpose of this study, data collected from 4 children presenting severe dyslexia and from 16 children who were not native French speakers were excluded. Addition data from 8 children whose recordings were unusable due to sampling issues were also excluded. A total of 295 children were included in the dataset. Table 4.1 presents the characteristics of the children for each grade (number, age, gender), together with their reading level expressed in the number of correct words read per minute (CWPM ; Maeder *et al.* 2018) and their non-verbal reasoning level measured with the PM38 test (Raven *et al.*, 2003). Both scores are congruent with the children's age according to established standards.

Oral readings of the same text by 20 adults were also recorded as references. These individuals were recruited in the lab on a voluntary basis. Sixteen of them have a university degree, four of them are graduate students. All adults were assessed as expert readers using a multidimensional fluency scale.

Material The participants were asked to read a text written for the specific purpose of the experiment and based on a short story for children (Friot, 2007). This is a 174 words narrative text with 13 declarative sentences each containing between 7 and 18 words. Among the 8 long sentences ($>$ 12words), 4 had no comma and 4 included commas to help distinguish phrases within the sentence. The average lexical frequencies of both long sentence types are equivalent (Lété et al., 2004), 4,528 occurrences for a million word ($SD = 325$) for the sentences with punctuation marks and 4,406 occurrences for a million words ($SD = 340$) for the others. The text is presented in Appendix 1.

Procedure The children were recorded on location at their respective schools where they were asked to read in a quiet room near their classrooms. The text they were asked to read aloud (described below) was presented to them at the time of each recording session in a counterbalanced order with another text used in another study. The adults were recorded in an experimental room in the lab where we followed the same procedure. All participants were recorded on their first reading of the text and were given no opportunity to prepare or rehearse.

We asked our participants to read "as if they were reading a story to a pre-schooler". This instruction was intended to (1) encourage an expressive reading and (2) to be ensure that it would be understood by the youngest readers. During each oral reading, we recorded participants' voice and breathing movements. Voice audio was recorded using a Schur Beta 53 headset microphone with a Behringer MIC100 amplifier. The breathing movements were recorded with two respiratory belts RESP150 placed on the thorax and the abdomen. Amplifier and belts signals were acquired with a Biopac MP150 to synchronise voice and breathing recordings. The headset microphone allowed to capture the breathing sound during the vocalization.

Table 4.1 – Characteristics of the participants according to their grade level

Grade	2	3	4	5	6	7	adults
Number	61	47	54	63	33	37	20
Females	27	25	26	23	15	21	10
Mean age	7 : 11(3.8)	8 : 11(3.4)	9 : 11(5.1)	10 : 11(4.5)	11 : 11(4.8)	12 : 8(5.4)	29 : 6(85)
Scope	(7 : 2 - 8 : 5)	(8 : 5 - 9 : 5)	(8 : 5 - 10 : 5)	(10 : 3 - 12)	(11 : 4 - 13)	(11 : 8 - 13 : 9)	(21 : 58)
CWPM	58(22)	80(31)	100(30)	108(42)	116(30)	143(76)	NA
PM38 raw score	33.6(9)	37.5(8)	39.7(6.8)	43.7(6.8)	38.6(5.9)	41.5(6.3)	NA

Note : means (standard deviations). Age is expressed in years : months. CWPM = number of words correctly read per minute.

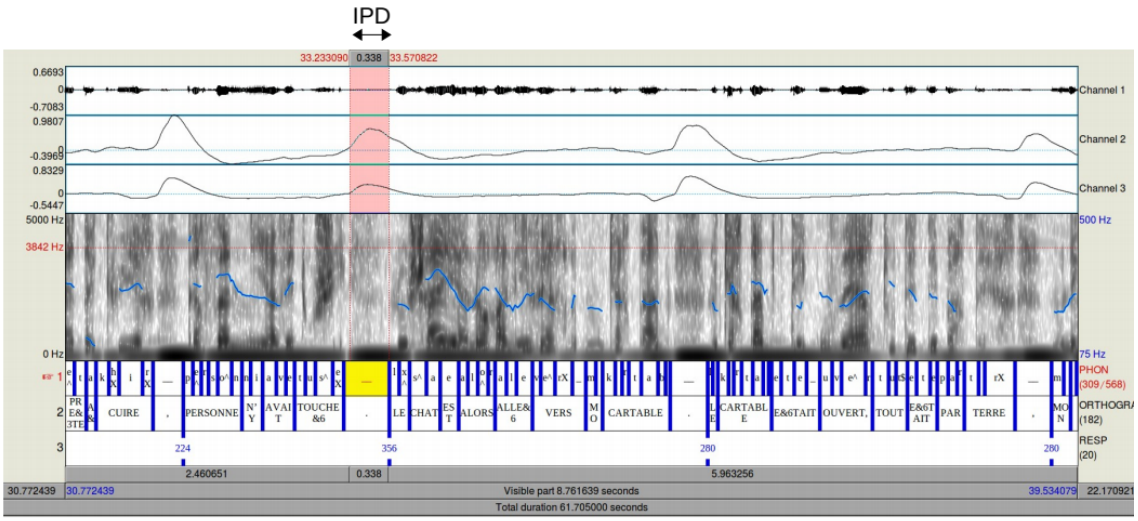


FIGURE 4.1 – Praat screenshot showing audio signal on channel one, respiratory belt signals on channel 2 (abdominal) and 3 (thoracic) and annotations of phonemes and pauses, words and inhalation to phonation delay (IPD). IPD was measured manually using audio and belts signals

Data Preparation The audio signals were automatically aligned to the text with a statistical model with phonetic triphone models, pronunciation dictionary and trigram model. As the children's' readings contained numerous pronunciation errors, omissions, hesitations and occasional repetition, each alignment was then manually corrected using Praat (Boersma, 2020). At the end of this process, the pronunciation database contained 1,241 correct and 1,111 incorrect pronunciations of the 174 words contained in the text, including 67 correct and 15 incorrect pronunciations comprising internal pauses, some of which were breathing pauses.

The pauses were annotated as either breathing or non-breathing pauses, using the audio clue of air intakes and the respiratory belts signals. For breathing pauses we also annotated the inhalation to phonation delay (IPD), as shown in Figure 4.1. Lastly, we computed 16,895 pauses from the 315 total readings, among which 8,900 were annotated as breathing pauses.

Data Analyses For each participant, various parameters concerning pauses duration (mean duration, within-participant variability) and pauses frequency, number of pauses per word of text (all pauses, ungrammatical pauses, pauses occurring at punctuation marks and pauses in long sentences with and without intra-sentential punctuation marks) were extracted and computed. After standard global analyses of all pause durations and all pause frequencies, we focused on breathing pauses (frequency, duration and IPD).

All parameters were retrieved and analysed using R (Team, 2019). The duration data were log-transformed before analysis (Campione et Véronis, 2002). We tested the effect of grade level using hierarchical linear mixed effect model with lmer function (lme4 R package, Bates *et al.* 2015). As the participants were from 14 different classrooms, the classroom itself was introduced as a random effect of level 2. This random effect of classroom was significant for every parameter tested (intra class correlation $> .05$), except for one. For each parameter, we compared the empty model and the model with the grade level as a fixed factor, using ANOVA. We then reported the significance of grade level with $\Delta\chi^2$(ddl) and p-value. When the grade level had a significant effect

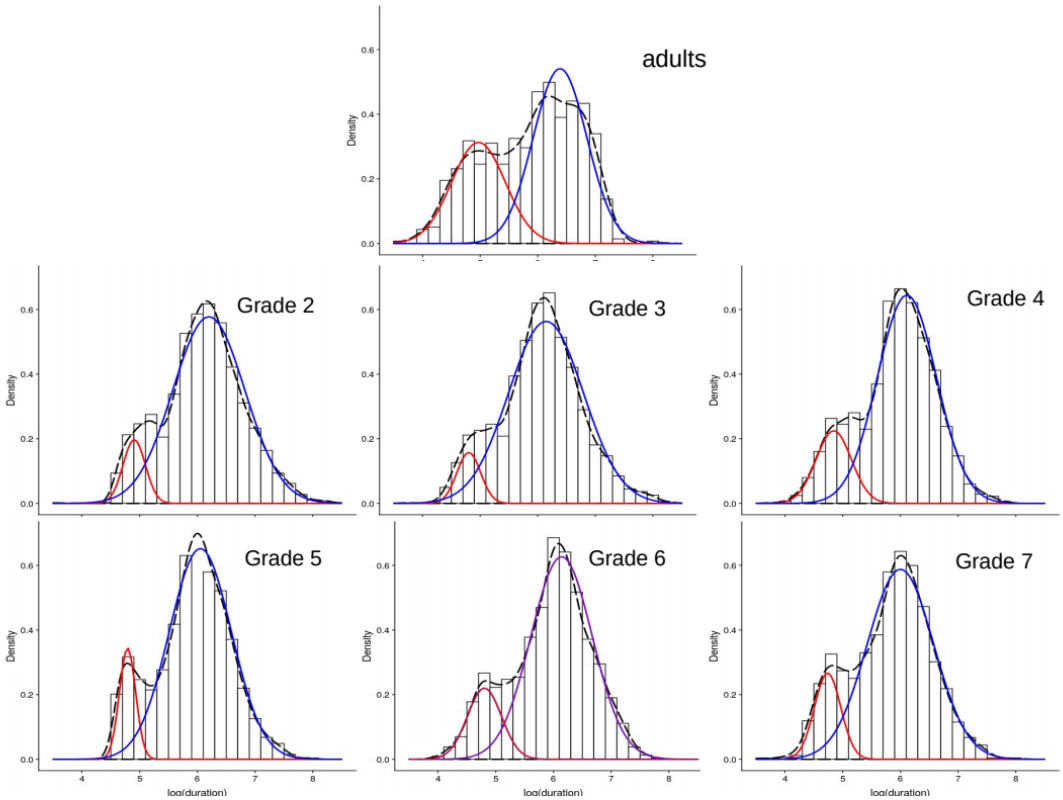


FIGURE 4.2 – Distribution of log duration pauses for each grade level : histograms and density curves (dashed line). Distributions have been approximated using a mixture model of two gaussians (short pauses = red line, long pauses = blue line)

on the parameter, we tested the significance of differences between grade level side by side, using the Tukey post-hoc test. We then reported only the highest p-value among all tests performed. Finally, we reported mean and standard deviation for each group and each parameter. When there was no significant difference between groups, we gathered the groups and reported the mean and the standard deviation of the cluster.

Results

Pause duration Figure 4.2 shows the distribution of log duration of pauses for each level. A Shapiro-Wilk test showed that, at each level, pauses duration is significantly different from the normality ($p < .001$ for each level). In accordance with previous studies (Campione *et al.*, 2002; Goldman *et al.*, 2010; Bailly et Gouvernayre, 2012), the distribution appeared bimodal at every level. We approximated the bimodal distributions using Gaussian Mixture Model with the normalmixEM function of the mixtools package (Benaglia *et al.*, 2009). The models converged for each level. The parameters calculated for each approximation are presented in Table 4.2 and plotted on Figure 4.2.

As our data were fitted with bimodal distributions as recommended by Campione et Véronis (2002), a distinction was made between short pauses (SP) centerd around the first gaussian (on average around 125ms) and long pauses (LP) centerd around the second gaussian (on average around 470ms). For the next stage of our analysis we drew a threshold line between SP and LP. As shown in Table 4.2, the means of gaussians are different from on grade to another. To take

TABLE 4.2 – For each grade, parameters of Gaussian mixture models for log duration pauses

	Grade													
	2		3		4		5		6		7		adults	
	SP	LP	SP	LP	SP	LP	SP	LP	SP	LP	SP	LP	SP	LP
λ	.096	.904	.081	.919	.173	.827	.121	.879	.163	.837	.152	.847	.366	.633
μ	4.91	6.20	4.80	6.15	4.84	6.11	4.79	6.05	4.80	6.14	4.74	6.00	4.97	6.38
σ	.193	.625	.206	.651	.306	.512	.141	.538	.295	.533	.228	.575	.467	.466
M(ms)	135	492	122	466	126	450	120	424	121	464	114	403	144	589
T(ms)	220		190		220		200		200		180		220	

Note : λ, μ, σ, computed using normalmixEM function of the R mixtool package ; with corresponding mean in ms (M) and threshold between short and long pauses (T) corresponding to the pass of the Gaussian mixture.

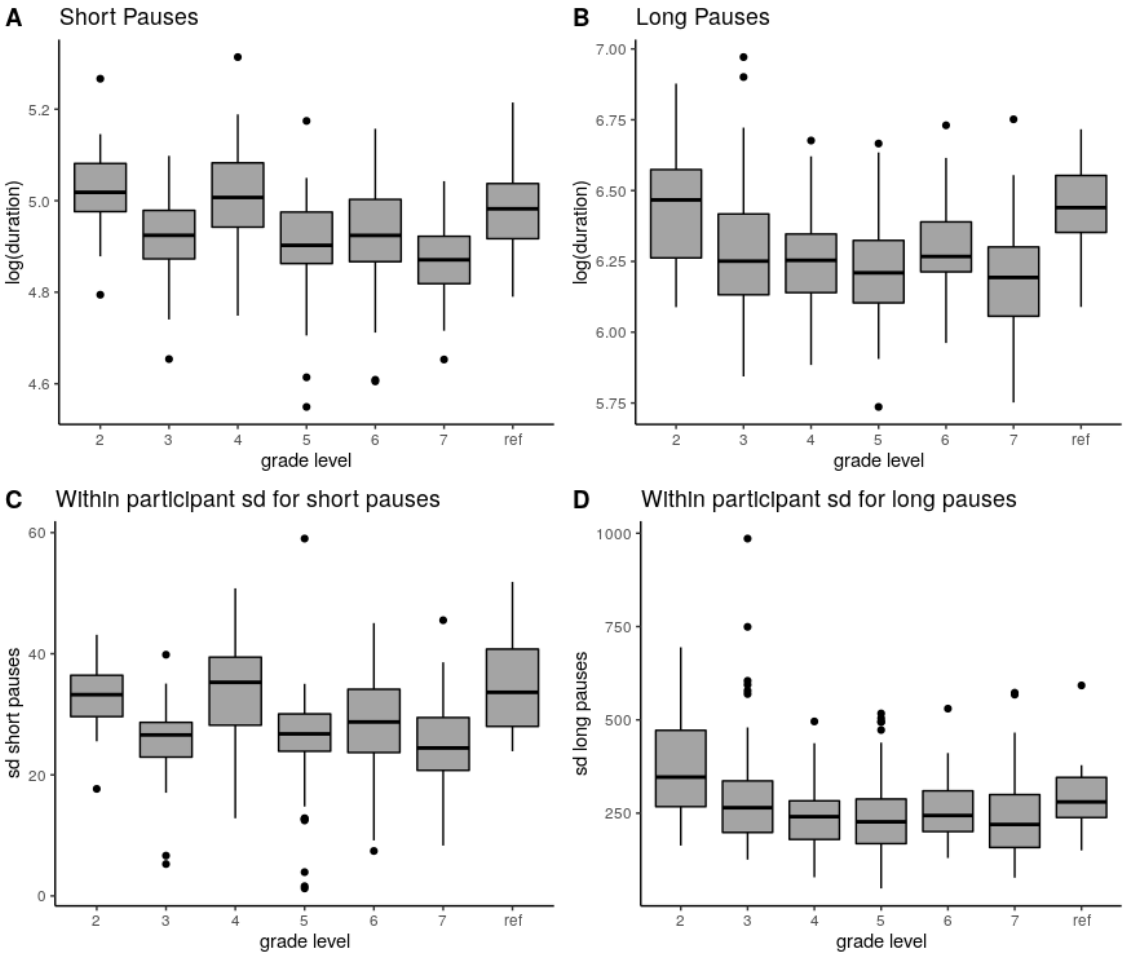


FIGURE 4.3 – Boxplots of pause durations corresponding to grade level, for short pauses (SP) in A and long pauses (LP) in B. Evolution of within-participant variability (standard deviation) according to grade level for SP in C and LP in D

that difference into account and to precisely classify pauses, the threshold between SP and LP was defined separately for each grade, as the distribution minimum between the two gaussians. Thresholds for each grade are presented in Table 4.2, together with the mean duration of each type of pause. Our results are congruent with those of Campione et Véronis (2002) and Bailly et Gouvernayre (2012).

We then looked at the effect of the grade level on pause duration. The grade level had a significant effect on both SP ($\Delta\chi^2(6) = 40.5$, $p < .001$) and LP ($\Delta\chi^2(6) = 28.7$, $p < .001$) mean duration. Overall, pause duration tended to decrease with grade and to be longer for adults than for the oldest children (see Figure 4.3A and 4.3B). More precisely, the Tukey test on SP showed that 2nd graders, 4th graders and adults produced significantly longer SP ($M = 151$ms, $SD = 14$) than the 3rd, 5th and 6th graders ($M = 136$ms, $SD = 13$; $p < .001$). Seventh graders produced significantly shorter SPs ($M = 130$ms, $SD = 10$) than any other group ($p < .029$). The Tukey test on LP revealed that 2nd graders and adults had significantly longer LPs ($M = 630$ms, $SD = 117$) than children in grades 3 to 7 (from grade 3 to 7 : $M = 528$ms, $SD = 103$; $p < .001$).

The effect of grade level is also significant on the within-participant variability (i.e., standard deviation). For SP duration (see Figure 4.3 C ; $\Delta\chi^2(6) = 37.5$, $p < .001$), the grade effect seemed to follow the same pattern for variability than for duration : 2nd graders, 4th graders and adults

produced less regular SPs than children in in higher grades. The grade level also had a significant effect on within-participant variability of LP duration (see Figure 4.3 ; $\Delta\chi^2(6) = 30.9$, $p < .001$), essentially due to the high variability of LP duration observed among second graders ($M = 373$ms, $SD = 135$) compared to other grade levels and adults ($p < .024$). This effect might be explained by the fact that the youngest readers made numerous hesitation pauses. Older children produced more regular LPs. It is also worth noting that the adults presented significantly more variation in their LP duration ($M = 293$, $SD = 94$) than children between grades 4 and 7 ($M = 244$, $SD = 102$; $p < 0.04$).

Pause frequency We specifically considered numbers of SPs and LPs per word of the text, i.e. pause frequency, as a barometer for bimodal distribution of pauses duration. Grade level showed to have a significant effect on the frequency of both SPs ($\Delta\chi^2(6) = 28,79$, $p < .001$) and LPs ($\Delta\chi^(6) = 35.83$, $p < .001$). The overall amount of SPs and LPs decreased with grade level (see Figure 4.4A and 4.4B). More precisely, the Tukey tests showed that second graders produced SPs more frequently ($M = 0.09$, $SD = 0.04$) than all other participant ($p < .001$). From grade 3 to 7, there was no significant difference between grade levels ($M = 0.05$, $SD = 0.03$). The developmental pattern of LPs appeared to differ slightly. Second and third graders made significantly more LPs ($M = 0.36$, $SD = 0.20$) than both the older children and adults ($p < .001$). The frequency of LPs remained stable for students in grades 4 to 7 and adults ($M = 0.17$, $SD = 0.05$). It is worth noting here that variability of LP number among students in one grade was particularly high in grades 2 ($SD = 0.18$) and 3 ($SD = 0.21$), and decreased fairly significantly after grade 4 ($SD < 08$). This observation confirms the degree of variability of reading skills in the first year of reading learning. Grade level also had a significant effect on the ratio of SPs versus LPs (see Figure 4.4C ; $\Delta\chi^2(6) = 29.79$, $p < .001$). Adults made significantly more short pauses ($M = 0.37$, $SD = 0.17$) than children ($p < .014$).

Ungrammatical pauses (UP) are interesting in that they indicate a lack of planning and anticipation, or decoding difficulties. Speakers tend to place this type of pause, for example, between a noun and its specifier or adjective, after a preposition or in a word. Our results showed that grade level was strongly correlated to UP frequency (see Figure 4.4D ; $\Delta\chi^2(6) = 33.5$, $p < .001$). The frequency of UPs made by second and third graders ($M = 0.18, SD = 0.15$) was roughly three times higher ($p < .001$) than in the other children. These pauses then dramatically decreased in fourth graders and stabilize in children in grades 5 to 7 ($M = 0.06$, $SD = 0.05$; $p < .003$).In adults, UP's frequency is close to zero. Importantly, adults did produce UPs. Indeed, our recordings of the participants' first readings of the text revealed that decoding difficulties did occur for some adults. We also looked at the ratio of ungrammatical pauses among all the pauses (see Figure 4.4E). Grade level had a significant impact on this ratio ($\Delta\chi^2(6) = 37.7$, $p < .001$). Although we expected the ratio to decrease from grades 2 to grade 4 and then stabilize with the acquisition of automaticity, our data showed a significantly higher ratio in third graders ($M = .44$, $SD = .18$) compared to second graders ($M = .32$, $SD = .13$; $p < .001$). The ratio decreased in fourth graders ($M = .29$, $SD = 14$; $p < .001$) and appeared to stabilize in children from grades 5 to 7 ($M = .20$, $SD = .13$). However, it remained higher than the ratio we observed in adults ($M = .06$, $SD = .03$; $p < .001$). What stood out was that the ratio was not yet adult-like in seventh graders. Variability in UP number between participants was particularly high in second and third graders (respectively, $SD = .13$ and $.18$). Once again, this could be due to the variability of reading skills during the first years of reading acquisition.

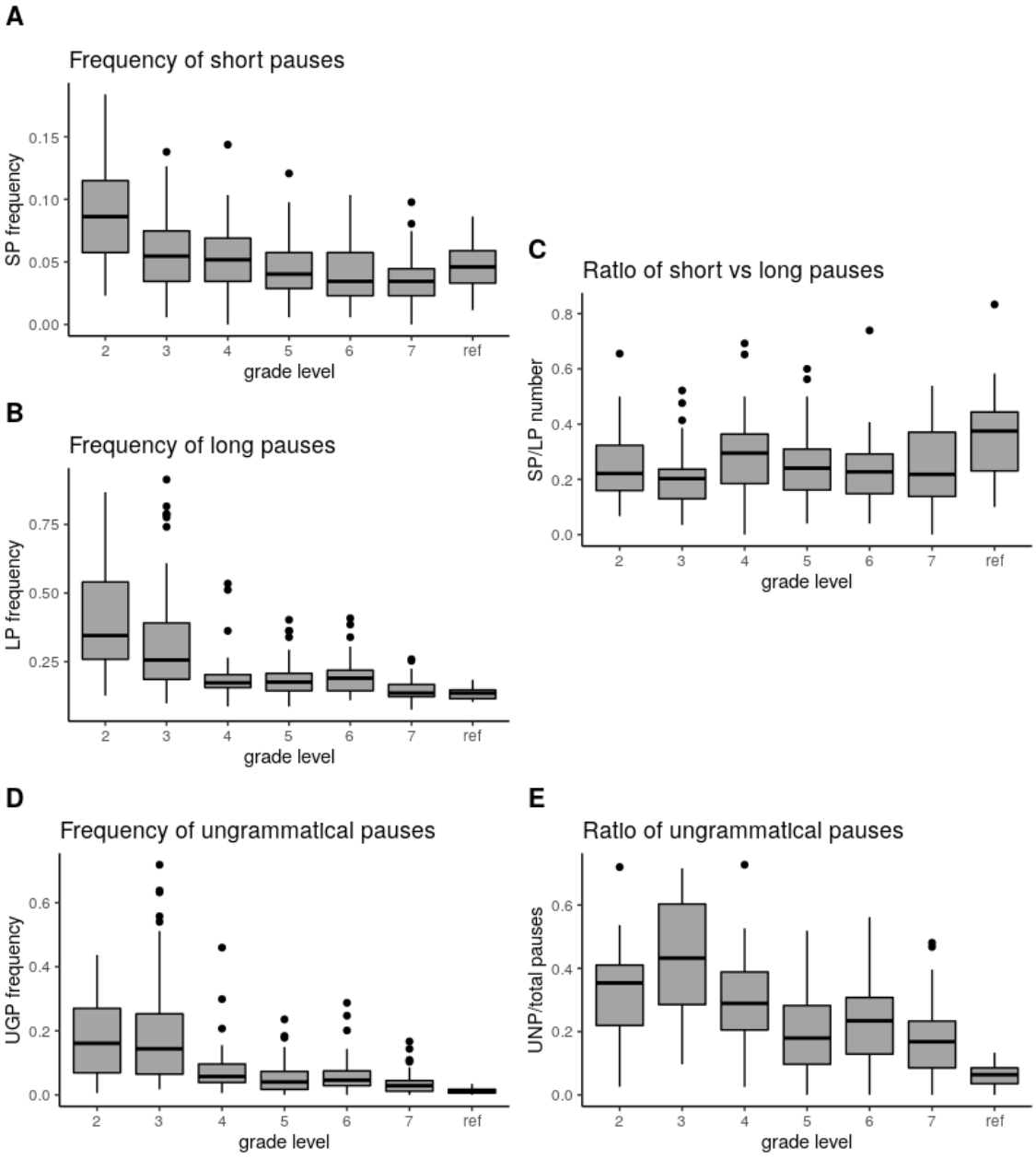


FIGURE 4.4 – Boxplots of pauses frequency according to grade level : frequency of short pauses (A), long pauses (B) and ungrammatical pauses (D) according to grade level, ratio of short and long pauses (C), ratio of ungrammatical pauses (E) according to grade level.

To sum up, SP and LP frequency decreased and seemed to level off in children from grades 4. The number of LPs produced by seventh graders is adult-like but they produced fewer SPs than the adults in this study. Our results show that ungrammatical pauses occurred often in the initial years of reading acquisition and were used more frequently by seventh graders than the adult readers. Pause planning seemed fairly efficient in children from grades 4.

Breathing pauses (BPs) Our first consideration here was looking at the proportion of BP among the short and long pauses. BPs are almost always LPs and, on average, represent roughly 70% of the total LPs. Among the 8,900 BPs analysed, only 15 were classified as SPs. We decided therefore to study BPs with no cleavage between short and long pauses. In the next stage of analysis, we considered BP as one type of pause (that might include a small proportion of SP). The grade level had no significant impact on the BP ratio (see Figure 4.5B ; $\Delta\chi^2(6) = 9.95$, $p = .12$) which was around 58% ($SD = .13$) of the overall number of pauses. Grade level, however, appeared to have a significant effect on BP frequency ($\Delta\chi^2(6) = 37.6$, $p < .001$). BPs decreased rather dramatically in children from grades 2 to grade 7 and adults (see Figure 4.5A). Grade 2 children made significantly more BPs ($M = 0.26$, $SD = 0.10$) that children from grade 4 and adults ($p < .001$), which we also observed in the third graders ($M = 0.20$, $SD = 0.10$; $p < .05$). The number of BPs levelled off from grades 4 to 7 and adults ($M = 0.12$, $SD = 0.04$).

Grade level also had a significant effect on BP duration ($\Delta\chi^2(6) = 23.7$, $p < .001$). Given that BP represents 70% of the LP, it was not surprising when we observed the same pattern as for LP duration : BP duration for second graders and adults ($M = 715$, $SD = 160$) were significantly longer ($p < .022$) than BP duration in children from grades 3 to 7 ($M = 588$, $SD = 161$). To take a more precise look at BP patterns, we focused specifically on the inhalation to phonation delay (IPD). Grade level also showed to have a significant effect on IPD ($\Delta\chi^2(6) = 23.7$, $p < .001$). In second graders, IPD was significantly higher ($M = 529$, $SD = 117$; $p < .001$) than in children from the other grades and adults ($MD = 413$, $SD = 110$). This indicated to us that second graders were less competent at consciously planning their BPs and simply tended to breathe as soon as they finished talking instead of inhaling before they began. Adults always inhaled just before beginning to speak and exhibited normal duration IPD of around 400ms ($SD = 66$). Second graders made longer BPs, which appeared to be caused by longer IPDs, revealing a lack of planning. Adults produced IPDs similar to children from grades 3 to grade 7, but they made a longer pre-inspiratory pause, leading to a longer BP. We observed that children older than grade 4 appeared to have near adult-like planning of their BPs. As for LP duration, the IPD and BP internal duration variability in each subject is affected by grade level (IPD : $\Delta\chi^2(6) = 42.7$, $p < .001$; BP : $\Delta\chi^2(6) = 29.9$, $p < .001$). The pattern was the same for both internal variability : a high variability in second graders (IPD : $M = 299$, $SD = 128$; BP : $M = 376$, $SD = 158$), that extended to third graders (IPD : $M = 206$, $SD = 111$; BP : $M = 310$, $SD = 200$) and decreased in older children (IPD : $M = 159$, $SD = 80$; BP : $M = 237$, $SD = 101$), supporting a lack of anticipation in the first year of reading acquisition.

To assess breathing coordination and planning, we looked at the frequency of ungrammatical BP (UBP). Indeed, if the BPs are planned, they should fit the syntax and be grammatical. As expected, the grade level had an impact on the frequency of UBPs ($\Delta\chi^2(6) = 32.3$, $p < .001$). Children in grades 2 and 3 made significantly more UBP pauses ($M = 0.08$, $SD = 0.06$) than the others ($p < .001$). The number of UBP was divided by three in the case of fourth graders ($M = 0.03$, $SD = 0.04$; $p < .02$). It then levelled off in children from garde 5 to grade 7 ($M = 0.02$, $SD = 0.02$).

In adults, the number of UBPs was null, effectively confirming that BP planning appears fairly established in fourth graders, even if not completely adult-like, probably because of late decoding issues. Finally, to confirm this observation, we computed the ratio of ungrammatical pauses among BPs. Again, grade level showed to have a significant effect ($\Delta\chi^2(6) = 36.7$, $p < .001$). This ratio was significantly higher in children from grades 2 and 3 ($M = 29\%$, $SD = .16$) than for older children ($M = 18\%$ in fourth grade, $p < .001$). The ratio continued to decrease and levelled off at 9% for children from grades 5 to 7. For the adults, the ratio was null, expressing perfect coordination of breathing during oral reading.

In summary, BP patterns develop in parallel with reading skills (e.g., matching decreases of UP and IPD), and lung development (e.g., decrease of pauses number). Children seem to acquire breathing-reading coordination around the grade 4 level. However, while BP planning is observable starting at grade 4, this skill is still not adult-like by the age of grade 7.

Punctuation effect on pausing patterns As previously reported by Martin (2011), punctuation marks in written language are important signals that tell the reader when prosody is required. In this study, we examined how children use punctuation to calibrate their pauses and breathing when reading aloud. We also compared the occurrence of ungrammatical pauses in long sentences with and without intra-sentential punctuation marks. Our hypothesis was that children use punctuation marks as visual aids for planning pauses, and consequently, that young readers would be more likely to produce ungrammatical pauses while reading sentences with no punctuation.

To test this hypothesis, we began by looking at the number of pauses occurring at punctuation marks. The text we asked participants to read aloud included a total of 22 punctuation marks (9 commas and 13 periods). As shown in Figure 4.6A, nearly all punctuation marks were followed by a pause. Our results showed a close correlation between grade level (age, reading expertise) and the number of these pauses occurring at punctuation marks ($\Delta\chi^2(6) = 20.7$, $p = .002$). The adults in our study paused at every punctuation mark ($M = 21.6$, $SD = 0.9$) while the children readers skipped a few of these cues ($M = 19.5$, $SD = 2.7$) ($p < .001$), more commonly commas. We then took a closer look at the types of pauses occurring at punctuation marks (see Figures 4.6B and 4.6C). Most of these pauses at punctuation marks were LPs (81% on average) and BPs (77% on average), with SPs occurring only rarely Grade level also had a significant effect on the ratio of BPs among punctuation pauses ($\Delta\chi^2(6) = 20.7$, $p = .002$). However, no significant difference was observed in readers from grades 2 to 6, for whom 79% ($SD = .13$) used punctuation marks as cues to breathe. Seventh graders and adults, however, were less likely to breathe at punctuation marks ($M = 71\%$, $SD = .14$), as they took fewer breaths in general while reading aloud ($p < .02$). These results confirmed our initial proposal that punctuation serves as a useful aid to help young readers plan and coordinate their breathing while reading.

We then proceeded to compare the number of pauses in longer sentences both with and without commas and with the same lexical complexity (Lété *et al.*, 2004). To take into account the large difference in the total number of pauses made by readers ranging from second graders to adults, we computed the ratio of pauses occurring in sentences with commas versus in sentences without commas, (Figure 4.6D). Grade level had a significant effect on the ratio ($\Delta\chi^2(6) = 18.55$, $p = .005$). For second graders, a t-test showed a ratio over 1 ($M = 1.13$, $SD = 0.25$; $t(56) = 3.68$, $p < .001$, $d = 0.52$). Second graders tended to pause more frequently when reading sentences with no punctuation. For children from grades 3 to 6, the ratio decreased significantly ($p = .04$) and stabilized around 1. Upward from grade 7, this ratio continued to decrease ($p < .02$) and

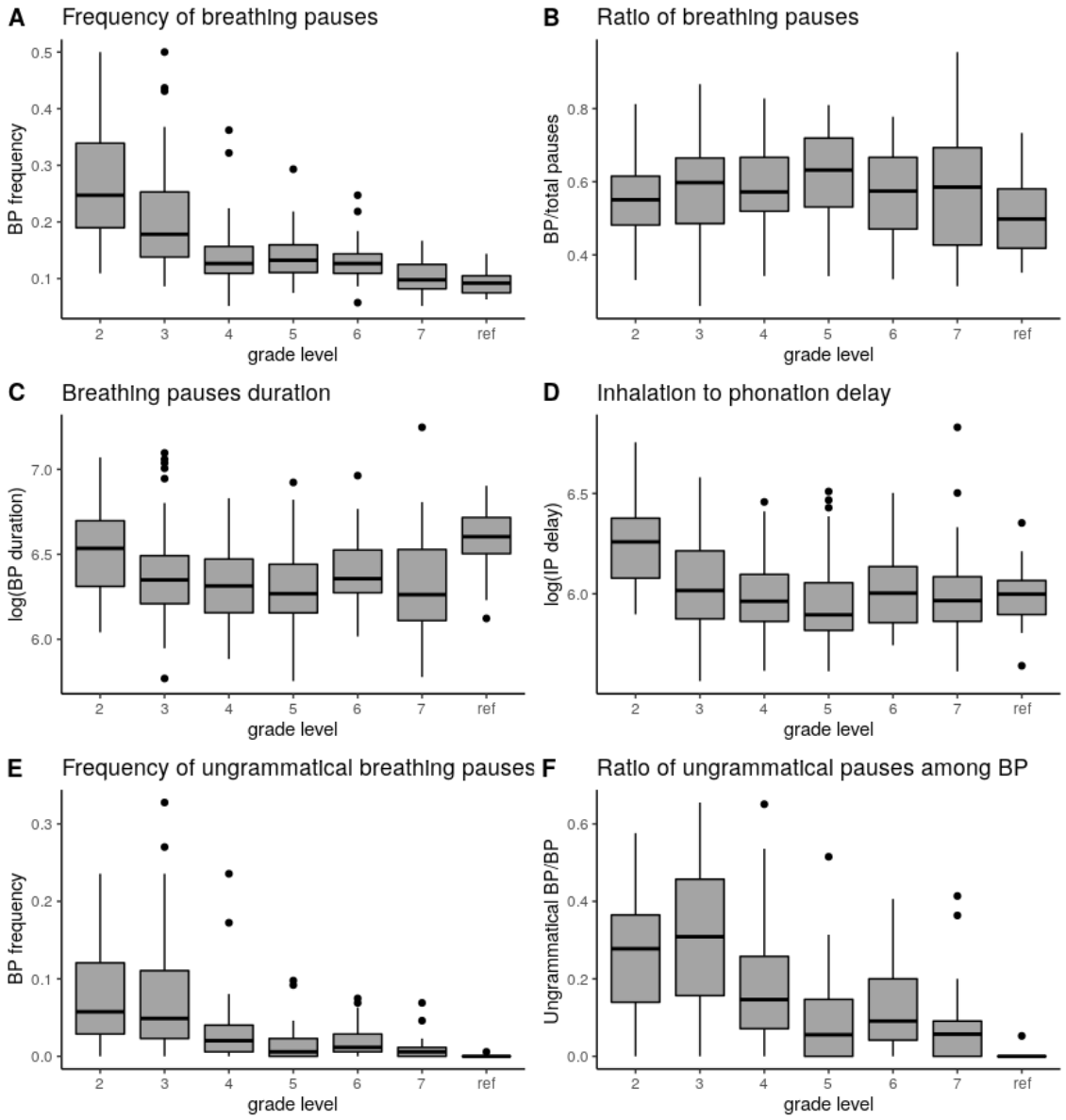


FIGURE 4.5 – Boxplots of breathing pauses from grades 2 to grade 7 and adults : frequency of breathing pauses (BP) (A), percentage of BP among the total pauses (B), BP duration with log transformation (C), inhalation to phonation delay (IP) after log transformation (D), frequency of ungrammatical BP (E) and ratio of ungrammatical pauses among BP (F)

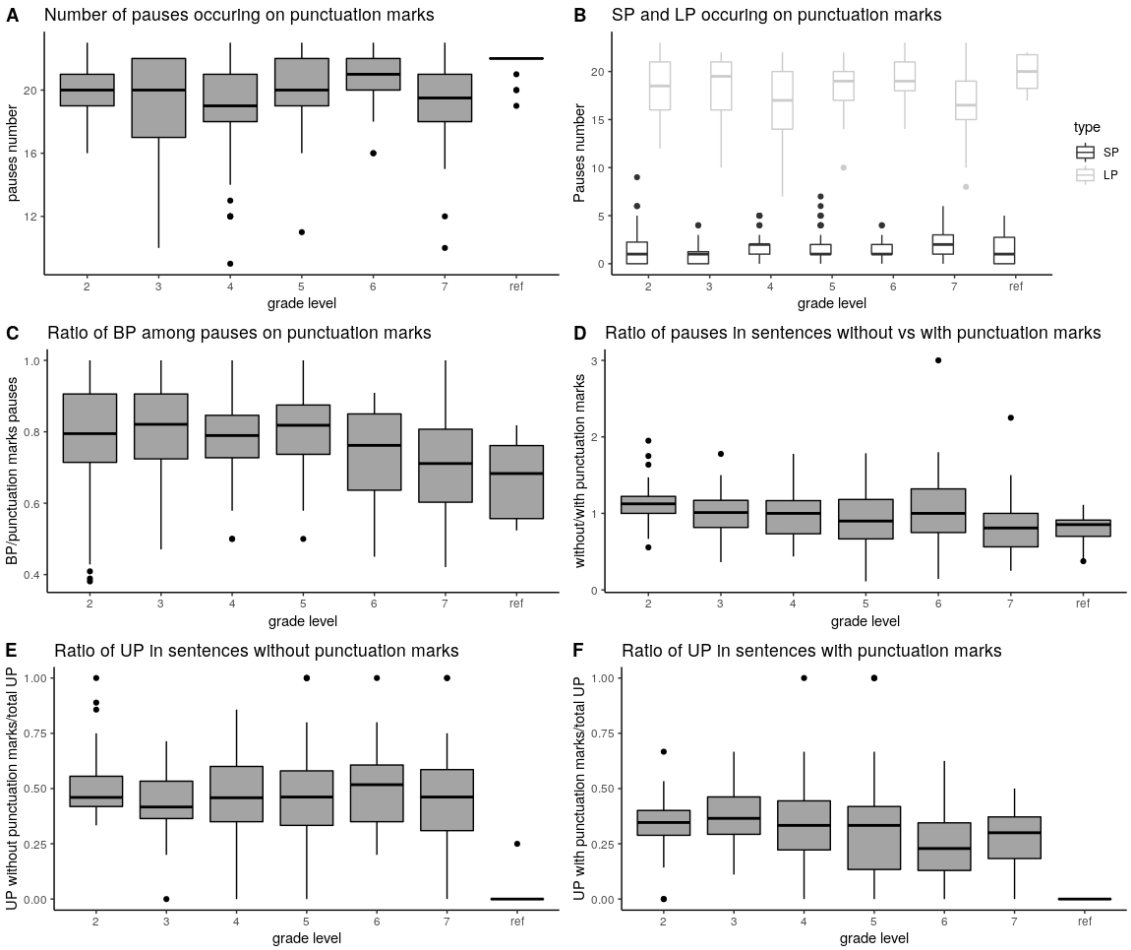


FIGURE 4.6 – Boxplots of pauses occurring at punctuation marks for readers in grades 2 to 7 and adults : number of pauses occurring at punctuation marks (A), number of short and long pauses occurring at punctuation marks (B), ratio of punctuation marks used to breathe (C), ratio of pauses in long sentences with and without punctuation marks (D), number of ungrammatical pauses (UP) occurring in long sentences with no punctuation marks (E) and ratio of UP occurring in sentences with punctuation marks (F).

became significantly smaller than 1 ($t(37) = -3.161$, $p = .003$, $d = 0.61$). For seventh graders and adults, pauses were more frequent in the presence of punctuation marks ($M = .79$, $SD = 35$). This observation effectively illustrates the inductive effect of punctuation on pausing for expert readers. For younger readers, even if they are not using all the punctuation marks to pause, punctuation may nonetheless help them to plan their pauses and avoid making ungrammatical pauses. Figures 4.6E and 6F show the ratio of ungrammatical pauses in sentences, with and without punctuation marks respectively. Due to the quasi-absence of ungrammatical pausing, adult ratios were not computed and were removed from the model testing the effect of grade level. This is the only model for which no random effects were associated with the class variable. Grade level had no impact on ungrammatical pausing in sentences without punctuation marks ($\Delta\chi^2(6) = 6.16$,$p = .401$). Forty-seven percent ($SD = .20$) of ungrammatical pauses occurred in long sentences with no help of punctuation marks. On the contrary, grade level had a significant effect on ungrammatical pausing in sentences that included punctuation marks ($\Delta\chi^2(6) = 37.6$, $p < .001$). For readers in primary grades, grades 2 to 5, 34% ($SD = .18$) of ungrammatical pauses occurred in sentences with punctuation marks. For middle school readers, grades 6 and 7, this ratio lowered significantly to 25% ($SD = .17$) ($p = .03$). The adult participants made no grammatical pauses for either type of sentence.

Taken together, these results tell us that punctuation marks are used for LPs, mostly BPs, and occur as early as grade 2. Grade level appeared to have only a negligible effect, likely due to lung maturation. However, intra-sentential punctuation marks clearly helped children to plan pauses. When children readers were successful with decoding, ungrammatical pauses occurred less frequently in the presence of punctuation marks. This effect was more pronounced in readers at the middle school level.

Discussion

A developmental pattern of pausing acquisition We observed that some aspects of pausing acquisition are present from the beginning of reading learning. This can be seen in bimodal pause duration distribution, breathing that almost always occurs during long pauses of which 58% are used for breathing, and punctuation marks representing 70% of long pauses that are largely used for breathing. We also found, however, that several variables developed as reading acquisition advanced. Figure 4.7 provides an overview of the evolution of the main parameters. We were able to identify four principal stages in pausing acquisition of French readers.

- **Grade 2** is characterized by numerous pauses, particularly ungrammatical pauses. As pausing patterns appear to be closely linked with word decoding skills, this stage involves no planning of either syntactic or respiratory pauses. Children at this age begin to rely on punctuation to help them identify appropriate pause placement, and have more difficulty in the absence of punctuation. The between-participant variability is high for most of the studied parameters, and reflect important differences in rates of individual reading acquisition. Similarly, within-participant variability reflecting unstable reading rate is high.
- **Grade 3** is characterized by overall fewer pauses and shorter pause duration. Coordinating linguistic content with speech breathing begins. Nevertheless, frequent ungrammatical pauses are still present at this stage and inter- and intra-variability remains elevated. Word decoding skills begin to improve, which leads to increased reading rates at the expense of pause planning.

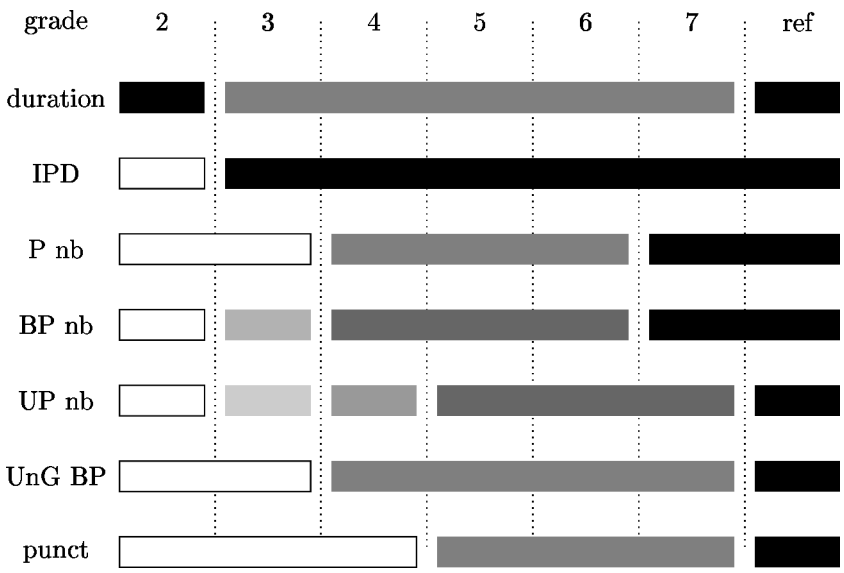


FIGURE 4.7 – Summary of the results on pausing acquisition in reading development. Black corresponds to adult-like values. Level of grey for children darkens for values closest to adults values. The parameters presented in this scheme are long pause duration, inhalation to phonation delay (IPD), number of pauses (P nb), breathing pauses (BP nb), ungrammatical pauses (UP nb), ungrammatical breathing pauses (UnG BP), and number of ungrammatical pauses occurring in long sentences with punctuation (punct).

- **Grades 4-6** is characterized by breathing maturity and planning of pauses. The overall frequency of pauses decreases considerably, including ungrammatical pauses. Children at this stage rely more on punctuation to plan their pauses when possible. Pause patterns appear to eventually stabilize after several years, but enough errors remain at this stage to still fall short of adult-like use of pausing
- **Grade 7** is characterized by pausing patterns that are nearly adult-like. Overall planning of breathing pauses seems to be efficient while a few ungrammatical pauses still occur, and the duration of pauses are still slightly shorter than adult pauses.

To summarize, proper use of pausing patterns appears to develop alongside other reading skills. The most crucial aspect lies in the ability to consciously control, or plan, where these pauses occur, particularly pauses that are used for breathing. Our study showed that planning acquisition begins to develop relatively slowly during the first year of reading acquisition. More noticeable progress really only begins at the 4th grade level when reading skills become more automatic and children have gained enough confidence to focus more on planning instead of decoding. Pausing acquisition, therefore, is still in developmental stages at the end of primary grade level and continues through middle school. By the time children reach grade 7, pausing patterns are generally quite adult-like.

Interpretation of the main results : relation between pauses, decoding and expressivity
We observed a bimodal distribution of pause duration in adults distinguished by use of either short or long pauses, as observed by Bailly et Gouvernayre (2012) and Campione et Véronis (2002). Clearly, short and long pauses are not linked to the same linguistic functions. Long pauses, that are more often used for breathing, develop in an interesting way. Indeed, children from grades 3 to 7 make shorter long pauses than second graders and adults. However, second graders' use of longer pauses during oral reading occurs for different reasons than it does for adults. The most

likely reason is that they are still learning to read and extended pauses represent difficulties or hesitation. Longer pauses simply give them time to think about decoding, which is to say they engage in a cognitive activity. As a consequence, the duration of pauses made by very young readers is linked directly to the level of their ability to decipher the next word, which can involve certain variability and might also lead to longer pauses ($>$ 1000ms).

Adults are far more likely to consciously plan their use of long pauses with certain linguistic purpose : marking major syntactic boundaries can make their reading expressive and enhance listener comprehension. Children, even in grade 7, have not yet mastered their patterns of pausing. Once decoding is automated (around grade 3), long pauses are used but are less pronounced than pauses used by adults. Consequently, the variability of long pause duration is higher in adults. Their syntactic pauses are planned and used to highlight the meaning of the text and achieve greater expressiveness. This could explain the increase of within-variability of pause duration in adults. Adult readers may play more with pauses duration to emphasize the significance of major syntactic boundaries versus minor ones which boosts the expressiveness of their reading. In the specific context of breathing pauses, we observed that the greatest difference between second graders and adults was essentially decomposition of the breathing pauses. Adults make a long pre-inspiratory pause and a short inhalation before speaking, while second graders make a short pre-inspiratory pause and a long inhalation and post-inspiration pause. Second graders tend to breathe when they finish talking, leading to long inhalation to phonation delay. Adults and older children are better at planning their breathing and tend to take a breath just before they begin speaking. In summary, long pause duration can mask two different strategies : a lack of planning and decoding issues as seen in 2nd graders and a will to emphasize major syntactic boundaries with long pauses for expert readers. Children from grades 3 to 7 seem to fall in between the two in that they are more adept at planning their breathing than second graders, but lack the use of emphasis more common in adult reading.

Another interesting observation is the increase of ungrammatical pausing for readers transitioning from 2nd to 3rd grade. Indeed, a previous study on children described a decrease of inappropriate pauses from grades 1 to 2 (Miller et Schwanenflugel, 2008) and 2 to 5 (Álvarez-Cañizo *et al.*, 2015). However, in spite of higher fluency skills1, we observed that 3rd graders made more ungrammatical pauses than 2nd graders. One possible explanation for this increase could be that 3rd grade readers focus more on speed and less on accuracy, despite the fact that they pause planning is still immature. Reading at a faster rate leads to less planning in pausing, and certainly breathing (Grosjean et Collins, 1979). Second graders, by contrast, are less confident decoding, read more slowly and tend to focus more on decoding, which gives them more time to plan their pauses. From grade 4, children seem to be more able to plan their pauses, both for breathing and to underline relevant syntactic boundaries, and ungrammatical pausing tends to decrease significantly. That said, ungrammatical pauses still occur for these readers, most likely due to decoding issues and lack of complex syntax anticipation. Moreover, beginning at the 4th grade level, punctuation begins to play a more important role in the pause planning process. Punctuation at this stage enables young readers to visualize syntactic boundaries, allowing them to plan where exactly they should place their pauses.

In terms of lung and coordination maturation, breathing coordination may not fully mature before age 10 (Wiechern *et al.*, 2018). We observed, however, a mature inhalation-to-phonation delay in readers as early as the 3rd grade. This suggests that children are able to plan their breathing patterns before they master syntactic pauses. Children in the 3rd grade, however, still

made a number of ungrammatical breathing pauses, which is more indicative of a lack of planning. To better understand this contradiction, we might consider the dynamics of pause duration. It is likely that 3rd graders reading at faster rates didn't plan their breathing pauses, but simply took breaths when they needed air, such that rapid inhalations were more closely linked to reading rate (Grosjean et Deschamps, 1975). The decrease in inhalation to phonation delay may instead be more closely associated to "rushed" reading, and less to the capacity for planning. In grade 4, ungrammatical breathing decreased significantly in accordance with a shorter inhalation-to-phonation delay, showing more mature planning of breathing pauses. This stage could indeed be related to the stage in breathing coordination observed by Hoit et al. (1990) around the age of 8. Even if speech coordination is not yet matured at this point, children rely to a certain degree on punctuation to better plan their breathing patterns while reading. The remaining ungrammatical breathing pauses that occur in the primary grades could be explained by remaining decoding difficulties, and are likely less to do with a lack of planning but more indicative of hesitation. Indeed, young readers stopping to decipher a word or hesitate will likely make longer pauses that can then be used to breathe during the corresponding cognitive activity.

Limitations of the study In this study, we chose to record and analyse the participants' first reading of the provided text. None of the participants, children and adults, were given the opportunity to read the text before reading it aloud. Our data reflect then the online treatment of the text's syntax and the anticipation in reading aloud. Our results would probably have been very different if the children had the occasion to silently read the text beforehand. Participants' individual reading levels would, however, likely have impacted the results by showing greater differences in planning between poor readers struggling with comprehension and subsequent planning difficulties, and more advanced readers able to understand and remember the text and experience no difficulty in planning their pauses reading a familiar text. It could potentially be interesting to consider the evolution of pauses between a first reading, as we did, and a second reading after a preliminary review of the text as preparation.

Additional types of pauses might also have been worth investigating. We did not annotate individual pause functions, apart from breathing (e.g., hesitation, syntactic, cognitive activity such as decoding or comprehension). Duration and number of pauses could potentially be linked to these pause functions. Similarly, we did not precisely identify those pauses occurring at a line break. These pauses in particular could indeed occur due to oculomotor functioning and would be less associated with conscious pause planning. Additionally, these kinds of pauses could occur more frequently in young children struggling to anticipate and process a quick line return. Examining this data with eye-tracking support could also provide valuable insight into pause planning. Finally we computed here the frequency of pauses related to the number of word of the text read. It could also be interesting to compute the frequency of pauses related to the amount of word pronounced. The development stages could then be different, mainly for young readers who tend to make a lot of hesitations and repetitions, leading to a number of word pronounced very different from the number of word printed.

Our group of adults were primarily teachers and PhD students and each expert reader. Our 7th grade participants were representative of their age category in terms of fluency. Consequently, the reading level of this group, as with the other groups of children, ranged from novice (struggling) to advanced (expert). This difference of mean reading level in the adult study population may explain the slight differences between grade 7 and adults. Some children are in fact adult-like

readers, while others may still struggle with reading fluency. Another fact might also have an impact on the reading levels of grade 6 and 7 students. In primary school, particularly in grades 2 and 3, reading aloud is a daily exercise. In our adult population, reading aloud is also frequently practiced. Children in French middle school are much less likely to be asked to read aloud in the context of a classroom activities. This type of exercise is then more exceptional for most students, more so than for the other participants. This difference could have an impact on their achievement, especially in their first reading of a text.

Finally, we purposefully confined our study to describing the pausing pattern. We did not consider the interplay or relationship between comprehension, fluency and pausing. The potential links between other reading skills and pausing acquisition could help to elaborate on the stages we observed in this study. For example, the pronounced within and between-participant variability in grades 2 and 3 may have resulted from differences in fluency and comprehension acquisition between children. In the same way, a longitudinal study from grades 2 to 5 could potentially be used for a more precise examination of how conscious planning of pauses develops, and could reveal more about specific relationships between the different analysed parameters.

Conclusions and perspectives

We designed our model to include several parameters of pausing (i.e., duration, number/frequency, correlation with breathing and punctuation marks) in order to characterize acquisition of expert reading pausing patterns through reading development. Our results largely correspond with findings from previous studies on adults and children, but we were also able to provide new insights on reading prosody. We observed that acquisition of pause planning in our participants happened in four stages, with a markedly rapid learning curve occurring during first years of reading learning before stabilizing toward the end of the primary grade levels and finally reaching maturity at grade 7. We also observed that punctuation marks play an important role in breathing planning and use of syntactic pauses. These observations could be particularly useful in terms of developing methods for teaching pausing planning during reading learning, for example, by emphasizing the importance of punctuation, or finding ways to help children extract syntactic structure to plan their pauses. When learning music, children are explicitly taught where and when to breathe or pause to best interpret a melody. They visualize where to pause and breathe at specific places on the music score. Introducing this type of structured, more systematic approach to teaching novice readers how to integrate pausing when reading aloud could help struggling readers with pause planning acquisition and reduce ungrammatical pauses that can be disruptive and prevent overall comprehension during oral reading.

Acknowledgements

We thank the school teachers and pupils for welcoming Erika Godde and Manon Metz and wearing respiratory belts. Anne-Claire Dugu recorded the adult readers. Margaux Manka, Anne-Laure Piat-Marchand and Camille Deshayes patiently corrected numerous phonetic alignment

Declaration

Funding This research has been conducted as part of “Fluence”, an e-FRAN project founded by the “Investissement d’Avenir” program handled by the French “Caisse des dépôts”.

Résultats majeurs **Étude 3**

On observe une bimodalité de la distribution de la durée des pauses, courtes et longues, dés le CE1.

Le développement de la planification des pauses est rapide du CE1 au CM1 avec une diminution du nombre et de la durée des pauses, du nombre de pauses non grammaticales et s'accompagne de la mise en place de la coordination de la respiration avecun appui sur la ponctuation.

Du CM1 en 6^{e}, les compétences se stabilisent mais sont encore différentes des adultes. Elles sont finalement équivalentes aux adultes à partir de la 5^{e}.

Les adultes experts et les CE1 présentent une durée moyenne de pauses plus longue que les autres enfants, particulièrement des pauses respiratoires. Chez les CE1, ces pauses plus longues sont en partie dues à un manque de planification de la respiration, tandis que chez les adultes elles sont dues à des pauses plus marquées aux frontières syntaxiques importantes.

4.2 Étude 4 : Marqueurs acoustiques de pauses et perception de la fluence

Dans cette étude, nous nous interessons au lien entre paramètres acoustiques et perception de la fluence. Quels paramètres infuencent notre façon d'évaluer les lectures à voix haute avec l'EMDF ? Peu d'études ont été menées sur le lien entre paramètres acoustiques et évaluations subjectives. Trois études ont été mentionnées dans le chapitre 1, celles de Cowie *et al.* (2002), Benjamin *et al.* (2013) et Young-Suk *et al.* (2020). Ces trois études montrent un lien direct entre notre perception de la prosodie en lecture et les marqueurs acoustiques de celle-ci. D'une part, le score subjectif de fluence est plutôt lié aux paramètres de pauses. Les variations de F0 sont liées aux scores d'expressivité dans deux des trois études. Pour Young-Suk *et al.* (2020), toutes les dimensions du score subjectif à la MDFS sont liées aux paramètres de pauses (fréquence et durée), tandis que les variations de F0 sont un paramètre à part entière dans le modèle de prosodie proposé. D'autre part, la vitesse de lecture a aussi un lien avec le score global de fluence. Il est important de noter que ces modèles montrent également des relations transverses entre tous les paramètres acoustiques et subjectifs, révélant un lien complexe. Dans le chapitre 3 et 4, nous avons évalué les lectures du corpus transversal en utilisant une échelle subjective (étude 1) et des paramètres acoustiques (étude 3). Dans cette quatrième étude, nous allons nous intéresser au lien entre les paramètres acoustiques que nous avons mesuré (pauses et respirations) et les scores attribués subjectivement aux différentes dimensions de fluence. Comment ces marqueurs acoustiques influencent-ils notre perception de la prosodie d'une lecture ? Conformément aux études citées précédemment, la modélisation effectuée dans cette étude devrait permettre de lier les paramètres de pauses (durée et nombre), de respiration et de vitesse de lecture aux scores subjectifs à l'EMDF.

4.2.1 Méthodologie

Données

Nous utilisons ici les données du corpus transversal présentées dans le chapitre 2. Seul le texte 1 est utilisé dans cette étude. Les modèles sont donc ici construits en utilisant 293 lectures du texte 1 par des sujets du CE1 à la 5^{e} et 20 lectures adultes du même texte. Les caractéristiques des sujets sont présentées dans le tableau 2.1.

La méthodologie d'évaluation subjective de la fluence des lectures est présentée dans le chapitre 3, étude 1. Les scores considérés ici sont les moyennes des scores donnés par les trois évaluateurs experts sur le texte 1 avec l'EMDF. Les 20 lectures des adultes de référence ont été évaluées lors de la même session, avec le même protocole que les lectures d'enfants. Nous considérons donc ici le score total de fluence à l'EMDF, et les scores aux quatre dimensions.

La méthodologie pour le traitement des données acoustiques est présentée dans le chapitre 4, étude 3. Tous les paramètres étudiés ne sont pas intégrés dans la modélisation. Seuls certains paramètres ont été conservés :

- nombre de pauses courtes
- nombre de pauses longues
- nombre de pauses non grammaticales

- nombre de pauses sur un signe de ponctuation
- durée moyenne des pauses longues
- durée DIP

Ces paramètres correspondent aux paramètres de pauses mis en lien avec les scores subjectifs dans la littérature : nombre et durée des pauses, nombres de pauses non grammaticales (cf chapitre 1). Nous y avons ajouté des paramètres spécifiques à notre étude : séparation des pauses en pauses courtes et pauses longues et ajout d'un paramètre respiratoire, le DIP. Le nombre et la durée des pauses respiratoires présentant une corrélation forte avec le nombre et la durée des pauses longues, seuls ces derniéres ont été conservées dans nos modèles. Finalement, nous controlons également le taux d'articulation, nombre de mots prononcés par minute, et le taux de phonation, rapport temps de phonation sur temps total de la lecture. Cela nous permet de ne pas confondre pauses et vitesse de parole, et de s'assurer que les prédicteurs liés aux pauses expliquent bien une part de variance propre et non un taux de phonation.

Analyses statistiques

Toutes les analyses statistiques ont été réalisées avec le logiciel R (Team, 2019). Pour déterminer le lien entre les scores subjectifs à l'EMDF et les paramètres acoustiques, nous avons utilisé un modèle général linéaire multiniveau avec la classe comme paramètre aléatoire de niveau 2. Cet effet aléatoire est toujours significatif (ICC > 0.05). Les paramètres utilisés dans le modèle sont les suivants :

- nombre de pauses courtes
- nombre de pauses longues
- nombre de pauses non grammaticales
- nombre de pauses sur un signe de ponctuation
- durée moyenne des pauses longues
- durée du DIP
- taux d'articulation
- taux de phonation

Cinq modèles ont été réalisés : un modèle pour le score global de fluence et un modèle pour chacune des dimensions de l'EMDF : expressivité, phrasé, décodage et vitesse. Tous ces scores ont été ramenés entre 0 et 1. Une transformation logarithmique a été appliquée aux durées. Nous avons ensuite réalisé des régressions bêta en utilisant la fonction glmmadmb du package *glmmADMB* (Fournier *et al.*, 2012). Nous avons réalisé des régressions multiples, en comparant par une anova les modèles successifs et en conservant les paramètres les plus significatifs à chaque étape, avant de calculer les coefficients de régression et les coefficients de détermination pour chaque variable significative du modèle. Nous rapportons la part de variance propre des variables d'intérêt, c'est à à dire leur part de variance, une fois retirée la variance expliquée par les variables contrôles.

TABLE 4.3 – Comparaison des modèles multiniveaux pour le score global de fluence à l'EMDF

	Modèle 1	**Modèle 2**	**Modèle 3**
Effets fixes			
Constante	-1.050(0.17)***	-3.28(0.32)***	-2.18(0.38)***
Articulation	1.10(0.06)***	1.19(0.06)***	0.85(0.09)***
Ponctuation		1.74(0.27)***	1.67(0.25)***
PNG			-0.012(0.002)***
Effets aléatoires			
Niveau 2 (classe)	0.14(0.37)	0.11(0.33)	0.13(0.36)
Qualité de l'ajustement			
AIC	-621	-658	-681
adjR^2	.801	.822	.848
ΔadjR^2	-	.021	.026

Note : PNG : nombre de pauses non grammaticales
* p < .05, ** p < .01, *** p < .001

4.2.2 Résultats

Score global de fluence

Les résultats des régressions multiples pour le score global de fluence sont présentés dans le tableau 4.3. Une fois le taux d'articulation controlé, la régression permet de faire apparaitre les effets fixes simples du nombre de pauses sur des signes de ponctuation et du nombre de pauses non grammaticales. Le score de fluence augmente avec le nombre de pauses liées à la ponctuation, il diminue quand le nombre de pauses non grammaticales augmente. Le nombre de pauses liées à la ponctuation et le nombre de pauses non-grammaticales expliquent également une part propre de variance de respectivement 2.1 et 2.6 %. L'évaluation du score global de fluence est donc lié à la vitesse de lecture, via le taux d'articulation, mais également aux pauses, et particulièrement à leur fonction syntaxique, via le respect de la ponctuation et de la syntaxe. Finalement, le modèle obtenu explique 84.8 % de la variance du score global de fluence, ce qui en fait un bon modèle.

Dimension Expressivité

Les résultats des régressions multiples pour le score à la dimension expressivité sont présentés dans le tableau 4.4. Une fois le taux d'articulation controlé, la régression permet de faire apparaitre les effets fixes simples du nombre de pauses sur des signes de ponctuation et du nombre de pauses non grammaticales. Le score à la dimension expressivité augmente avec le nombre de pauses liées à la ponctuation, il diminue quand le nombre de pauses non grammaticales augmente. Le nombre de pauses non grammaticales est le prédicteur le plus important avec 19 % de variance du score à la dimension expressivité expliquée. Le nombre de pauses liées à la ponctuation explique également une part propre de variance de 5.2 %. L'évaluation du score à la dimension expressivité est donc essentiellement liée au respect de la syntaxe et de la ponctuation dans le placement des pauses. La vitesse de lecture n'intervient que de façon plus marginale. Finalement, le modèle obtenu explique 59.3 % de la variance du score à la dimension expressivité. C'est un modèle assez bon. En effet,

Table 4.4 – Comparaison des modèles multiniveaux pour le score à la dimension expressivité de l'EMDF

	Modèle 1	**Modèle 2**	**Modèle 3**
Effets fixes			
Constante	-1.40(0.24)***	-0.05(0.41)	-3.08(0.60)***
Articulation	0.89(0.09)***	0.46(0.14)	0.65(0.13)***
PNG		-0.019(0.005)***	-0.016(0.004)***
Ponctuation			2.85(0.43)***
Effets aléatoires			
Niveau 2 (classe)	0.14(0.37)	0.14(0.37)	0.10(0.31)
Qualité de l'ajustement			
AIC	-243	-258	-297
adjR^2	.485	.521	.593
ΔadjR^2	-	.036	.072

Note : PNG : nombre de pauses non grammaticales
* p < .05, ** p < .01, *** p < .001

nous n'avons utilisé que des variables liées au rythme dans les régressions. Les variables liées aux variations de F0, qu'on pourrait attendre dans la prédiction d'un score d'expressivité, n'ont pas été prise en compte ici.

Dimension Phrasé

Les résultats des régressions multiples pour le score à la dimension phrasé sont présentés dans le tableau 4.5. Une fois le taux d'articulation controlé, la régression permet de faire apparaitre les effets fixes simples du nombre de pauses liées aux signes de ponctuation, du nombre de pauses longues et de leur durée moyenne. Le score à la dimension phrasé augmente avec le nombre de pauses liées à la ponctuation et la durée moyenne des pauses, il diminue avec le nombre de pauses longues. Les nombres de pauses longues et liées à la ponctuation expliquent une part propre de variance de respectivement 3.1 et 1.4 %. La durée des pauses n'est que très faiblement explicative avec seulement 0.1% de variance expliquée. L'évaluation du score à la dimension phrasé est donc essentiellement liée à la vitesse de lecture et aux nombres de pauses. L'apport du respect de la syntaxe n'est que marginal. Finalement, le modèle final obtenu explique 77.6 % de la variance du score à la dimension phrasé.

Dimension Décodage

Les résultats des régressions multiples pour le score à la dimension décodage sont présentés dans le tableau 4.6. Une fois le taux d'articulation controlé, la régression permet de faire apparaitre l'effet fixe simple du nombre de pauses liées aux signes de ponctuation. Le score à la dimension décodage augmente avec le nombre de pauses liées à la ponctuation. Le respect de la ponctuation est présent mais n'est que très faiblement explicatif avec seulement 0.1% de variance expliquée. L'évaluation du score à la dimension décodage est donc essentiellement liée à la vitesse de lecture. Le modèle final obtenu explique 70.6 % de la variance de ce score. Ce modèle est donc relativement bon, bien que nous n'avons pas inclus de variables liées aux erreurs de lecture dans le modèle.

TABLE 4.5 – Comparaison des modèles multiniveaux pour le score à la dimension phrasé de l'EMDF

	Modèle 1	**Modèle 2**	**Modèle 3**	**Modèle 4**
Effets fixes				
Constante	-1.74(0.23)***	-3.93(0.46)***	-2.69(0.55)***	-7.51(1.82)***
Articulation	1.26(0.09)***	1.36(0.09)***	0.95(0.14)***	1.11(0.15)***
Ponctuation		2.16(0.40)***	2.47(0.39)***	2.28(0.40)***
Pauses longues			-0.012(0.003)***	-0.011(0.003)***
Durée des pauses				0.73(0.26)**
Effets aléatoires				
Niveau 2 (classe)	0.16(0.40)	0.13(0.35)	0.14(0.38)	0.12(0.35)
Qualité de l'ajustement				
AIC	-598	-623	-636	-642
adjR^2	.728	.742	.775	.776
ΔadjR^2	-	.014	.031	.001

Note : * p < .05, ** p < .01, *** p < .001

Table 4.6 – Comparaison des modèles multiniveaux pour le score à la dimension décodage de l'EMDF

	Modèle 1	**modèle 2**
Effets fixes		
Constante	-1.44(0.21)***	-2.40(0.42)***
Articulation	1.08(0.08)***	1.13(0.08)***
Ponctuation		0.94(0.37)**
Effets aléatoires		
Niveau 2 (classe)	0.13(0.36)	
Qualité de l'ajustement		
AIC	-559	-563
adjR^2	.703	.704
ΔadjR^2	-	.001

Note : PNG : nombre de pauses non grammaticales
* p < .05, ** p < .01, *** p < .001

Dimension Vitesse

Les résultats des régressions multiples pour le score à la dimension vitesse sont présentés dans le tableau 4.7. Une fois le taux d'articulation controlé, la régression permet de faire apparaitre les effets fixes simples du DIP, du nombres de pauses longues et de leur durée moyenne et du nombre de pauses liées à à la ponctuation. Le score à la dimension vitesse augmente avec le nombre de pauses liées à la ponctuation et la durée moyenne des pauses, il diminue avec le nombre de pauses longues et le DIP. Les autres prédicteurs, bien qu'améliorant le modèle, n'explique pas de variance du score de vitesse. L'évaluation du score à la dimension vitesse est donc essentiellement liée à la vitesse de lecture et aux nombres de pauses. La gestion des pauses ne semble intervenir que marginalement. Le modèle final obtenu explique 77.5 % de la variance du score à la dimension vitesse.

4.2.3 Discussion

Au delà du taux d'articulation, et donc de la vitesse, les différentes variables liées aux pauses apportent une participation propre aux scores des différentes dimensions de la fluence. Le respect de la ponctuation, présent dans les modèles obtenus pour chaque dimension, semble être une variable importante dans l'appréciation subjective de la fluence, aussi bien dans les dimensions prosodiques, phrasé et expressivité, que dans les dimensions d'automaticité, décodage et vitesse. Le nombre et la durée des pauses sont des facteurs qui ont finalement un impact assez faible sur l'évaluation subjective. Le respect de la ponctuation, et de la syntaxe pour l'expressivité, semble influencer plus fortement les scores attribués.

Il est intéressant de remarquer que, comme dans l'étude de Young-Suk *et al.* (2020), les prédicteurs liés aux pauses sont présents dans les modèles de toutes les dimensions du score de fluence, y compris dans les dimensions d'automaticité. Le respect de la ponctuation est notamment retrouvé dans tous les modèles. On peut s'étonner de l'influence des variables de pauses dans l'automaticité.

TABLE 4.7 – Comparaison des modèles multiniveaux pour le score à la dimension vitesse de l'EMDF

	Modèle 1	**Modèle 2**	**Modèle 3**	**Modèle 4**	**Modèle 5**
Effets fixes					
Contante	-1.81(0.24)***	6.07(1.96)**	2.95(2.14)	4.33(2.24)	4(2.21)
Articulation	1.52(0.10)***	1.23(0.13)***	1.28(0.12)***	0.94(0.17)***	0.88(0.18)***
DIP		-1.19(0.29)***	-2.16(0.43)***	-2.16(0.43)***	-1.93(0.44)***
Durée des pauses			1.41(0.47)**	1.38(0.47)**	1.08(0.48)*
Pauses longues				-0.0088(0.0035)*	-0.011(0.003)**
Ponctuation					1.17(0.44)**
Effets aléatoires					
Niveau 2 (classe)	0.091(0.30)	0.11(0.33)	0.08(0.28)	0.07(0.26)	
Qualité de l'ajustement					
AIC	-926	-940	-948	-952	-957
adjR^2	.775	.77	.77	.77	.77
ΔadjR^2-	-	-	-	-	

Note : * p < .05, ** p < .01, *** p < .001

Plusieurs hypothèses peuvent expliquer cet apport. Les pauses sont liées à l'activité cognitive, particulièrement chez les jeunes lecteurs : ceux-ci font des pauses longues pour se donner le temps de décoder. Cependant dans notre modèle, le score de la dimension décodage est plutôt lié au respect de la ponctuation. On peut également penser que cet aspect de la lecture influence notre perception de la précision de décodage. Cette influence est cohérente avec l'intercorrélation des dimensions de la fluence mentionnée dans la littérature (Cowie *et al.*, 2002; Benjamin *et al.*, 2013). La prosodie parait donc influencer la perception du décodage et de la vitesse de lecture. Ainsi les juges ne pourraient juger indépendamment chaque dimension mais seraient influencés par une perception globale de la lecture, apparemment lié au respect de la ponctuation.

Deux particularités dans ces modèles attirent l'attention. D'une part dans le score de phrasé, le nombre de pauses non grammaticales, c'est à dire le respect de la syntaxe, n'apparait pas dans les prédicteurs, seul le respect de la ponctuation est prédicteur de ce score. Le nombre de pauses non grammaticales étant corrélé au taux d'articulatiion, il est possible que dans le jugement d'une minute de lecture, les juges soient plus sensibles aux frontières syntaxiques majeures, qui sont marquées par la ponctuation. D'autre part, la durée des pauses semble très peu influencer le jugement subjectif de la prosodie. En effet, la durée des pauses n'intervient que très marginalement dans le phrasé et la vitesse. De plus, le coefficient positif de la variable de durée des pauses dans le modèle de phrasé implique que le score augmente avec la durée des pauses. Cela peut paraitre étonnant, les pauses les plus longues étant, chez les plus jeunes, plutôt des marqueurs d'hésitations et donc de scores faibles. Il est possible que son implication soit négative chez les plus jeunes lecteurs et positive chez les plus grands. La régression linéaire a peut-être tendance à donner un effet global faible. En testant des modèles non linéaires, l'effet de la durée des pauses serait peut-être plus important. Cela appuierait l'idée que le sens de la durées des pauses diffère selon l'âge et les compétences des lecteurs. Concernant le phrasé, on pourrait associer ces pauses longues à un appui plus marqué de la structure syntaxique. Il est plus difficile d'interpréter ce coefficient pour le score de vitesse. En effet, comme l'ont montré Grosjean et Collins (1979) et Breznitz (1990) quand la vitesse de lecture augmente, la durée des pauses diminue. Cependant, ces observations concernent respectivement des adultes et des enfants accélèrant leur rythme de lecture au-delà de leur rythme normal. Nous nous intéressons dans cette étude à des enfants pour qui les compétences de lecture sont encore en cours d'acquisition et pour lesquels la consigne était une consigne de lecture pour autrui ("comme si tu lisais une histoire à des petits") et non de vitesse. Les scores faibles correspondent à une vitesse faible et les score élevés à une vitesse normale et non accélérée. Ces résultats relèvent donc probablement d'un mécanisme différent.

Finalement, comme dans les chapitres précédents, on peut remarquer la particularité de l'expressivité par rapport aux autres dimensions de la fluence. Tout d'abord le placement des pauses joue un rôle important dans la prédiction du score d'expressivité. Les respects de la syntaxe et de la ponctuation prédisent une part non négligeable du score. Contrairement aux autres dimensions pour lesquelles le taux de phonation explique une part très importante de variance, il intervient très peu dans le score d'expressivité. Il n'est donc pas nécessaire de lire vite pour être expressif, mais il semble nécessaire d'avoir un phrasé correct. Ces observations vont dans le sens de la littérature, avec une interdépendance des dimensions prosodiques de la lecture (Cowie *et al.*, 2002; Benjamin *et al.*, 2013; Young-Suk *et al.*, 2020). De plus, le modèle correspondant aux scores d'expressivité est le modèle le plus faiblement explicatif de toutes les dimensions. Ceci peut s'expliquer simplement par le fait qu'aucune variable liée aux variations de F0 n'a été introduite. En effet, la dimension d'expressivité dans l'EMDF est évaluée par la capacité à s'approcher du langage parlé et donc à

utiliser une large palette de variations mélodiques. Il est donc nécessaire d'intégrer de nouveaux prédicteurs pour améliorer le modèle du score d'expressivité.

Pour conclure, cette étude a permis de mettre en avant l'importance du respect de la syntaxe et de la ponctuation dans l'évaluation subjective de la prosodie. Cependant afin d'améliorer la prédiction des scores subjectifs, il serait nécessaire de prendre en compte les variations de F0 pendant la lecture. On se heurte alors à la question de l'objectif à atteindre. En effet, si le placement des pauses est très lié à la syntaxe du texte, l'expressivité dépend de l'interprétation que le lecteur fait du texte et est donc très variable d'un lecteur à l'autre. De ce fait, son évaluation objective reste problématique.

Résultats majeurs	**Étude 4**
Le respect de la ponctuation influence le jugement subjectif de toutes les dimensions de l'EMDF.	
Le nombre et de la durée des pauses n'ont que peu d'impact et le nombre de pauses non grammaticales aucun impact sur le jugement de toutes les dimensions de l'EMDF.	
On observe une influence importante du respect de la syntaxe sur le jugement subjectif du phrasé.	
Le jugement de l'expressivité est lié au respect de la syntaxe et de la ponctuation.	

4.3 Étude 5 : Marqueur acoustiques des pauses et compréhension

L'étude 3 a permis de mettre en évidence l'acquisition de la planification des pauses et de la respiration en lecture. Cette planification se met essentiellement en place du CE1 au CM1, période durant laquelle la durée des pauses et le nombre des pauses grammaticales et non grammaticales diminuent. Ces compétences se développent conjointement aux autres compétences littéraires que sont l'automaticité et la compréhension. Plusieurs études, présentées dans le chapitre 1, ont pu mettre en évidence des liens entre des marqueurs liés aux pauses et la compréhension écrite chez les jeunes lecteurs, en espagnol et français (québecois) (Álvarez-Cañizo *et al.*, 2015; Arcand *et al.*, 2014). Ces deux études montrent un lien entre le nombres de pauses non grammaticales et les scores à des tests de compréhension écrite, chez des enfants de CE2 et 6^{e}. Cependant dans ces deux études, les variables liées aux pauses utilisées dans les modèles se limitent aux nombre et durée des pauses et au nombre de pauses non grammaticales. Nous disposons dans le travail mené ici de marqueurs plus fins avec lesquels travailler, comme les pauses sur ponctuation. Par ailleurs,dans le chapitre 3.2, nous avons pu mettre en évidence le lien entre compréhension écrite, à la fois globale et en-ligne, et la dimension de phrasé dans les lectures. Dans l'étude 4, nous avons également pu mettre en évidence le lien entre phrasé et nombre et durée des pauses longues et attention à la ponctuation. La pause apparaissant comme un marqueur du phrasé, il apparait donc probable que nous retrouvions ce lien entre organisation des pauses et compréhension écrite.

Dans cette étude, nous allons donc d'une part chercher à retrouver ce lien entre caractéristiques des pauses et compréhension écrite, décrit dans la littérature et esquissé dans le chapitre 3.2. Mais nous allons également pour cela introduire les variables caractéristiques du développement que nous avons mis en évidence dans l'étude 3 sur les marqueurs acoustiques liés aux pauses et à la respiration, notamment la différenciation pauses courtes et longues et le DIP. Nous nous attendons à retrouver un lien entre pauses non grammaticales et compréhension, mais également à enrichir ce lien avec de nouvelles variables respiratoires. Comme dans le chapitre précédent, nous nous intéresserons à la fois à la compréhension écrite globale, utilisée traditionnellement dans la littérature, mais aussi à la compréhension écrite en-ligne.

4.3.1 Méthodologie

Données expérimentales

Pauses Nous utilisons ici les données acoustiques récoltées et traitées dans l'étude précédente. Nous avons conservé les paramètres de pauses les plus caractéristiques du développement de la planification en nombre, durée, ponctuation, grammaticalité et respiration.

- nombre : pauses courtes, pauses longues, pauses non grammaticale, pauses sur ponctuation
- durée : pauses longues, DPI

Seules les données des enfants ont été conservées pour cette modélisation. Une transformation logarithmique est appliquée aux variables de durées.

Compréhension Comme dans le chapitre précédent, nous utiliserons ici deux types de compréhension : compréhension globale et compréhension en ligne. Ces tests sont présentés dans le chapitre 2. Le score de compréhension globale est le taux de réussite aux questions. Il est compris entre 0 et 1. Le score de compréhension en-ligne est calculé à partir de la différence entre le NMCLM dans le texte signifiant et dans le texte non-signifiant. Pour ramener ce score entre 0 et 1, il est divisé par le NMCLM dans le texte signifiant. Les scores moyens, écart-type, minimum et maximum pour chaque niveau sont présentés dans le tableau 3.11.

Variables contrôlées Pour expliquer la compréhension écrite, nous controlons également les scores de vocabulaire, automaticité (NMCLM sur texte signifiant), raisonnement non-verbal et compréhension orale. Ces tests sont présentés dans le chapitre 2.

Analyses

Toutes les analyses statistiques ont été réalisées avec le logiciel R (Team, 2019). Pour déterminer le lien entre pauses et compréhension nous avons utilisé un modèle général linéaire multiniveau avec la classe comme paramètre aléatoire de niveau 2. Nous avons procédé à des régressions successives en complétant notre modèle avec le paramètre le plus significatif à chaque étape de sa construction. Les paramètres utilisés dans le modèle sont les suivants :

- niveau
- nombre de pauses longues
- nombre de pauses courtes
- nombre de pauses sur la ponctuation
- nombre de pauses non grammaticales
- durée des pauses longues
- durée du DPI
- raisonnement non verbal
- NMCLM
- vocabulaire
- compréhension orale
- compréhension écrite globale
- compréhension écrite en-ligne

Les interactions des paramètres les plus explicatifs avec le niveau et le NMCLM ont également été ajoutées dans les régressions. Deux modèles ont été réalisés : un modèle pour la compréhension globale, un modèle pour la compréhension en-ligne. Pour les scores de compréhension, nous avons effectué la transformation proposée par Smithson et Verkuilen (2006) afin de prendre en compte les possibles zéro pouvant poser problème dans le modèle. Nous avons ensuite réalisé les régressions bêta en utilisant la fonction glmmadmb du package *glmmADMB* (Fournier *et al.*, 2012). Nous avons réalisé des régressions multiples, en comparant par une anova les modèles successifs et en conservant le paramètre le plus significatif à chaque étape, avant de calculer les coefficients de régression et les coefficients de détermination pour chaque variable significative du modèle. Nous rapportons la part de variance propre des variables d'intérêt, c'est à à dire leur part de variance, une fois retirée la variance expliquée par les variables contrôles.

Table 4.8 – Comparaison des modèles multiniveaux pour le score de CE global

	Modèle 1	**Modèle 2**
Effets fixes		
Constante	-2.17(0.52)***	-1.25(0.65)
Vocabulaire	0.12(0.04)***	0.096(0.038)*
Niveau 3	-0.01(0.75)	-0.16(0.77)
Niveau 4	1.01(0.74)	0.44(0.79)
Niveau 5	0.45(0.81)	-0.04(0.83)
Niveau 6	-0.68(0.91)	-1.32(0.94)
Niveau 7	-0.36(0.84)	-1.03(0.88)
Voc*niv3	-0.016(0.047)	-0.010(0.049)
Voc*niv4	-0.11(0.04)**	-0.091(0.046)*
Voc*niv5	-0.10(0.04)*	-0.013(0.051)
Voc*niv6	0.039(0.05)	-0.013(0.051)
Voc*niv7	-0.049(0.045)	-0.023(0.047)
RNV	0.025(0.009)**	0.026(0.009)**
CO	0.071(0.029)*	0.067(0.029)*
PL		-0.007(0.003)*
Effets aléatoires		
Niveau 2 (classe)	3.8e-9(6e-5)	0.0013(0.036)
Qualité de l'ajustement		
AIC	-171	-175
$adjR^2$	.348	.369
$\Delta adjR^2$	-	.021

Note : RNV : raisonnement non-verbal, PL : nombre de pauses longues, CO : compréhension orale, Voc : vocabulaire

4.3.2 Résultats

Compréhension écrite globale

Les résultats des régressions multiples sont présentés dans le tableau 4.8.

Une fois controlés le vocabulaire, le raisonnement non-verbal et la compréhension orale, la régression multiple sur le score de CE globale fait apparaitre l'effet fixe simple significatif du nombre de pauses longues Le nombre de pauses longues apporte une contribution propre de 2,1 % à la variance du score de CE globale. Le score de compréhension diminue avec le nombre de pauses longues.

Nous avons ensuite évalué l'interaction des variables explicatives avec le niveau. Cette interaction est significative uniquement pour le vocabulaire ($p < .001$). Il n'y a aucune interaction significative du nombre de pauses longues ni avec le niveau, ni avec le NMCLM. On peut également noter que le modèle proposé ici est un modèle avec un coefficient de détermination assez faible ($R^2 = .369$). On peut finalement remarquer que la durée et la grammaticalité des pauses n'ont aucun d'impact sur ce score, seul le nombre de pauses longues a une incidence.

TABLE 4.9 – Comparaison des modèles multiniveaux pour le score de CE en-ligne

	Modèle 1	**Modèle 2**	**Modèle 3**
Effets fixes			
Constante	-2.17(0.18)***	-0.99(0.29)***	-9.7(0.3)***
Voc	0.077(0.011)***	0.068(0.01)***	0.067(0.011)***
NMCLM	-0.0042(0.0006)***	-0.0067(0.0008)***	-4.4e-3(1.1e-3)***
PL		-0.012(0.003)***	-2.5e-3(3.9e-3)
PL*NMCLM			-1.29e-4(4.1e-5)**
Effets aléatoires			
Niveau 2 (classe)	5.5e-8(2..3e-4)	2.6e-9(5.1e-5)	6.8e-9(8.2e-5)
Qualité de l'ajustement			
AIC	-492	-514	-520
$adjR^2$	.234	.305	.316
$\Delta adjR^2$	-	.071	.011

Note : NMCLM : nombre de mots correctement lus par minute, PL : nombre de pauses longues, Voc : vocabulaire

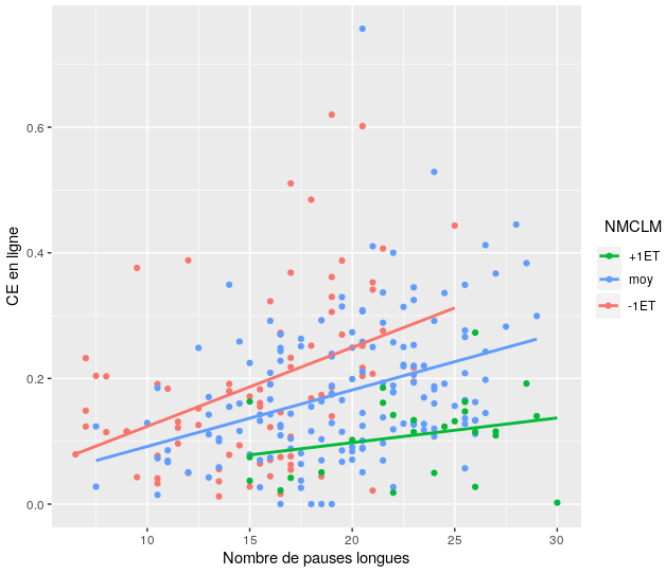


FIGURE 4.8 – Interaction nombres de pauses longues-NMCLM dans le score de compréhension écrite en-ligne. Score de compréhension écrite en ligne en fonction du nombre de pauses longues pour la valeur moyenne de NMCLM (bleu), plus un écart-type (vert) et moins un écart-type (rose).

Compréhension en ligne

Les résultats des régressions multiples sont présentés dans le tableau 4.9.

Une fois le vocabulaire et le NMCLM controlés, la régression multiple sur le score de CE en-ligne fait apparaitre l'effet fixe simple significatif du nombre de pauses longues. Le nombre de pauses longues apporte une contribution propre de 7.1% à la variance du score de CE en-ligne. Ce score diminue avec le nombres de pauses longues.

Nous avons ensuite évalué l'interaction des variables explicatives avec le niveau et le NMCLM. Cette interaction est significative uniquement entre le nombre de pauses et le NMCLM ($p<.001$) et explique 1.1% de variance. On ne constate aucune interaction avec le niveau scolaire. La figure 4.8 montre l'évolution des pentes du score de CE en-ligne en fonction du nombre de pauses longues pour la valeur de NMCLM moyenne et pour des valeurs à plus et moins un écart-type. On peut constater que l'impact du nombre de pauses est d'autant plus important que les enfants lisent lentement. On peut également noter que le modèle proposé ici est un modèle avec un coefficient de détermination assez faible ($R^2 = .316$).

4.3.3 Discussion

Comme attendu, les résultats de cette étude montrent bien un lien entre les pauses et la compréhension écrite, aussi bien globale que en-ligne, une fois le vocabulaire, raisonnement non-verbal, compréhension orale et NMCLM controlés. Cependant dans nos modèles, une seule variable de pause prédit négativement les scores de compréhension : le nombre de pauses longues. On ne retrouve ni le nombre de pauses non grammaticales, ni la durée des pauses ou de la respiration, ni l'utilisation de la ponctuation. L'absence de la respiration dans les prédicteurs peut-être lié au fait que 60% des pauses longues sont des pauses respiratoires. Il est plus étonnant de ne pas retrouver le nombre de pauses non grammaticales, comme Arcand *et al.* (2014). Notons que Benjamin *et al.*

(2013) observe également le nombre de pauses comme étant plus en lien avec le phrasé que le nombre de pauses non grammaticales. Ces modèles peuvent se rapprocher des modèles présentés dans l'étude 2 du chapitre 3. On retrouve en effet le nombre de pauses longues, là où apparaissait le score subjectif de phrasé.

Dans le modèle de compréhension en-ligne, on retrouve l'interaction avec le NMCLM. Cette fois, l'interaction concerne le nombre de pauses longues. On constate alors que, quand le NMCLM augmente, l'impact du nombre de pauses longues sur le score de compréhension écrite en-ligne diminue. Ainsi, l'impact du nombre de pauses longues est d'autant plus important quand les enfants lisent lentement. La compréhension des enfants lisant vite est moins impacté par les pauses longues. On peut interpréter cette interaction en faisant le lien entre pauses longues et pauses d'hésitation. Chez les lecteurs les plus faibles, les pauses longues sont souvent des pauses d'hésitations nuisant à la compréhension. Pour les lecteurs plus à l'aise, ayant acquis l'automaticité (donc avec un NMCLM plus élevé), les pauses longues peuvent être des marqueurs prosodiques visant à accentuer les frontières syntaxiques et donc facilitant la compréhension.

On constate également que le NMCLM n'intervient pas dans le modèle de compréhension globale, alors que le nombre de pauses longues intervient. Dans cette étude nous montrons donc, comme dans le chapitre précédent, que la prosodie est un marqueur important de la compréhension. La mesure du NMCLM n'étant pas forcément révélatrice de la compréhension globale d'un texte. Ces modèles restent cependant peu explicatifs.

Álvarez-Cañizo *et al.* (2015) constate aussi un effet significatif des variations de F0 sur la compréhension globale. Il serait donc intéressant dans un deuxième temps d'ajouter des variables liées aux variations de F0 dans ce modèle. Un modèle longitudinal permettrait également de cerner la directionnalité du lien prosodie-compréhension. En effet, les précédentes analyses de régression sont toutes effectuées sur des données transversales et ne permettent de mettre en évidence qu'une corrélation.

Résultats majeurs — **Étude 5**

Le nombre de pauses longues est négativement corrélé à la compréhension écrite globale et en-ligne et explique une part propre de leur variance.

Le NMCLM pondère l'impact du nombre de pauses longues dans la compréhension en ligne. Pour les enfants qui lisent lentement, un nombre de pauses longues élevé est plus révélateur d'une faible compréhension en-ligne, que pour les enfants lisant plus rapidement.

4.4 Conclusion

Les trois études présentées dans ce chapitre ont permis de mettre en évidence l'évolution des variables liées aux pauses dans le développement de la prosodie en lecture, leur impact sur la perception de l'auditeur et leur lien avec la compréhension du lecteur.

La planification des pauses s'acquiert de façon assez rapide dans les premières années d'apprentissage de la lecture, du CE1 au CM1. Il existe dés le CE1 une bimodalité de la distribution de la durée des pauses : pauses courtes et longues, comme chez les adultes. La coordination de la respiration avec la syntaxe du texte est acquise également rapidement, dés le CE2, en s'appuyant notamment sur la ponctuation, même si ces pauses restent plus nombreuses que chez les enfants plus âgés et les adultes. Les pauses respiratoires et syntaxiques restent cependant moins marqués chez les enfants que chez des adultes experts.

Le placement des pauses, c'est à dire le respect de la ponctuation et de la syntaxe, a un impact significatif sur le jugement subjectif de l'auditeur sur la lecture. Le nombre de pauses longues est, quant à lui, négativement corrélé à la compréhension globale et en-ligne. Cet effet est d'autant plus marqué que les enfants ont une automaticité faible.

Chapitre 5

Caractérisation du développement longitudinal de la prosodie par une évaluation automatisée

Dans les chapitres 3 et 4, nous avons abordé le développement de la prosodie en lecture et son lien avec la compréhension écrite par le biais d'une évaluation subjective, puis d'un point de vue acoustique, avec le cas des pauses. Nous n'avons cependant pas abordé les marqueurs acoustiques de l'expressivité car il n'existe pas une unique façon idéale de rendre un texte expressif. Dans ce dernier chapitre expérimental, nous proposons d'utiliser dans un même outil les deux méthodes d'évaluation abordées précédemment et d'y inclure la variabilité des productions possibles pour proposer un nouveau type de mesure de la prosodie en lecture. Cet outil sera utilisé ensuite pour évaluer les lectures de l'étude longitudinale afin d'explorer la causalité du lien entre prosodie en lecture et compréhension, encore peu documenté. L'étude 6 présente cet outil de prédiction automatique du score à l'EMDF prenant en compte la multiplicité des productions possibles à partir des marqueurs acoustiques et verbaux. Nous utiliserons ensuite l'outil présenté pour explorer le développement des différentes dimensions de la fluence des sujets suivis en longitudinal. Pour finir, les scores obtenus avec l'outil présenté dans l'étude 6 sont utilisés dans l'étude 7 pour proposer un modèle longitudinal d'acquisition de la prosodie et de la compréhension en lecture.

5.1 Étude 6 : Développement longitudinal de la prosodie mesurée par évaluation automatisée de la fluence

L'étude 1 présente l'évolution des scores à chaque dimension de l'EMDF pour les enfants du corpus transversal. Dans cette nouvelle étude nous souhaitons observer l'évolution de ces scores de façon longitudinal. Cependant, pour évaluer les lectures du corpus longitudinal, nous utilisons ici un outil d'évaluation automatique des dimensions de la fluence développé au GIPSA-Lab par Gérard Bailly à partir des données présentées dans cette thèse. Cet outil utilise à la fois les paramètres verbaux et prosodiques des lectures, et permet de comparer les lectures d'enfants à plusieurs lectures d'adultes experts. A partir de ces paramètres, l'outil attribue un score à chaque dimension

de l'EMDF pour la lecture évaluée. Ces scores sont ensuite utilisés pour construire des modèles de croissance pour chacune des dimensions. Nous observerons ici l'effet du niveau scolaire, mais aussi du texte lu, sur le développement de chacune des dimensions calculées par l'outil d'analyse automatique.

5.1.1 Présentation de l'outil d'évaluation automatique

Le travail présenté dans ce paragraphe fait l'objet d'un article en préparation, écrit avec Gérard Bailly, Anne-Laure Piat-Marchand et Marie-Line Bosse. L'outil présenté ici a été développé au GIPSA-Lab par Gérard Bailly. Ma contribution à son élaboration est la récolte et le traitement des données audios et des évaluations subjectives, et la validation de l'outil. Dans ce paragraphe nous résumons l'article en préparation, en présentant l'outil automatique, ses principes et sa validation.

Comme nous l'avons vu dans les chapitres précédents, le besoin d'outils d'évaluation de la fluence incluant la prosodie va grandissant, à la fois pour la recherche et pour l'éducation, notamment pour le suivi dans le temps des compétences en lecture des élèves. La difficulté majeure est la complexité de cette évaluation. Il existe en effet de multiple façons différentes de lire un texte à haute voix en ajoutant de l'expressivité. Le jugement de la qualité de l'expressivité dépend du lecteur, mais également de l'évaluateur. A l'heure actuelle, on trouve traditionnellement les deux types d'outils d'évaluation abordés dans les chapitre 3 et 4 : les échelles subjectives et les analyses acoustiques manuelles. Dans la littérature les analyses acoustiques se concentrent sur les marqueurs rythmiques, tels que les pauses. Les paramètres mélodiques et d'intensité sont peu exploités (cf chapitre 1). Quand ils le sont, c'est en comparaison avec une lecture experte unique. Ces analyses se heurtent donc à la problématique de la variété des productions acceptables. De plus, elles nécessitent un alignement précis de la voix et du texte, ce qui prend du temps dans le cas de lecture d'enfants contenant répétitions, hésitations ,... Bien que précises, ces analyses sont complexes à réaliser et utiliser. Dans une optique de simplification, des outils d'évaluation automatisée commencent à être développés. Un premier outil proposé par Bolanos *et al.* (2013) utilise à la fois des paramètres verbaux (NMCLM, taux de phonation, taux d'articulation et nombre de répétition) et acoustiques (localisation de l'accentuation, durée des mots et des syllabes, pauses silencieuses et non silencieuses et lien avec la ponctuation) pour prédire des scores d'enfant de CP au CE2 à l'échelle unidimensionnelle NAEP (Pinnell *et al.*, 1995). Pour cela, les auteurs utilisent un apprentissage supervisé (machine à vecteur de support). Les lectures sont également évaluées à l'aide de cette échelle par deux juges experts. Les auteurs obtiennent de bons résultats avec un coefficient de corrélation calculé/jugé de 0.86. Ils concluent cependant sur la difficulté de l'évaluation de l'expressivité, dû à un manque d'objectifs simples à atteindre. On peut également noter qu'ils utilisent une échelle unidimensionnelle et ne séparent donc pas les dimensions de prosodie et d'automaticité. Ce sont ces deux derniers aspects que nous proposons de prendre en compte dans l'outil d'évaluation automatique présenté ici en utilisant une échelle multidimensionnelle, l'EMDF et des références multiples. Une première version de notre outil avait été utilisée pour évaluer des lectures d'enfant de CE2 sur le texte "L'Alouette" (Godde *et al.*, 2017b). La méthode était alors la même que celle que nous présentons ici mais le texte choisi, "L'Alouette" (Lefavrais, 2005), destiné à évaluer le NMCLM, était trop complexe pour le public étudié. De plus la consigne et le texte en lui-même n'induisaient pas suffisamment de prosodie pour pouvoir être correctement évalué dans la dimension expressivité. C'est pourquoi nous présentons ici l'outil actualisé, utilisant des textes adaptés à l'âge des sujets étudiés, avec une consigne de lecture expressive.

Données audios

Les données audios utilisées ici proviennent des corpus transversal pour les textes 1 et 2 et longitudinal pour le texte 1 uniquement. Toutes les lectures sont utilisées soit un total de 744 lectures de 326 enfants et 20 adultes. Le protocole de recueil des données, les descriptions des cohortes et des textes sont disponibles dans le chapitre 2.

Evaluation subjective

Les 744 lectures ont été évaluées subjectivement en utilisant l'EMDF. L'utilisation de l'échelle et le protocole suivi sont présentés dans le chapitre 3.2. Toutes les évaluations réalisées ont été conservées pour cette étude. Au total, vingt-sept juges ont évalué 744 lectures. Nous disposons donc de 6853 évaluations à analyser. On peut noter que le texte 1, utilisé dans l'étude 1 pour déterminer l'effet du type d'évaluateurs sur les scores à l'EMDF, comporte beaucoup plus d'évaluations que le texte 2.

Evaluation objective

L'évaluation objective des lectures se base sur des paramètres verbaux et des schémas prosodiques. Afin de procéder à cette évaluation, les enregistrements sont alignés automatiquement sur les textes en utilisant un dictionnaire régulièrement mis à jour avec les nouvelles lectures et prenant en compte les différentes prononciations correctes et incorrectes de tous les mots des textes.

Caractérisation des paramètres verbaux Au delà du classique NMCLM, d'autres paramètres suceptibles d'influer sur les scores de décodage et vitesse ont été pris en compte :

- nombre de mots incorrects par minute (NMI)
- nombre de répétitions par minute (NR)
- nombre de pause intra-mots par minute (NPIM)
- nombre de noyaux vocaliques par minute (NV)

Pour les 438 lectures des 159 mots du texte 1, nous avons relevé 856 prononciations correctes et 1385 prononciations incorrectes. Cela correspond en moyenne à 5.38 prononciations correctes et 8.71 prononciations incorrectes par mot du texte. Pour les 306 lectures des 155 mots du texte 2, nous avons relevé 660 prononciations correctes et 988 incorrectes, soit 4.26 prononciations correctes par mot du texte et 6.37 prononciations incorrectes. On peut ainsi noter que les deux textes semblent assez difficile à décoder, le texte 1 plus que le texte 2.

Caractérisation des schémas prosodiques Le défi de la caractérisation des schémas prosodiques est de tenir compte de la multiplicité des productions acceptables. Hirst *et al.* (1998) proposent d'évaluer la prosodie des productions de synthétiseurs vocaux en utilisant une comparaison avec plusieurs références naturelles. Aprés plusieurs alignements (durée des segments, F0 et intensité), ils utilisaient alors comme mesure de la qualité prosodique la valeur moyenne quadratique de la distance entre la production de synthèse et la production naturelle la plus proche. Cette approche ne permet cependant pas de prendre en compte la globalité des productions possibles en se limitant à peu de dimensions et à une production cible. Le choix s'est porté ici sur la méthode

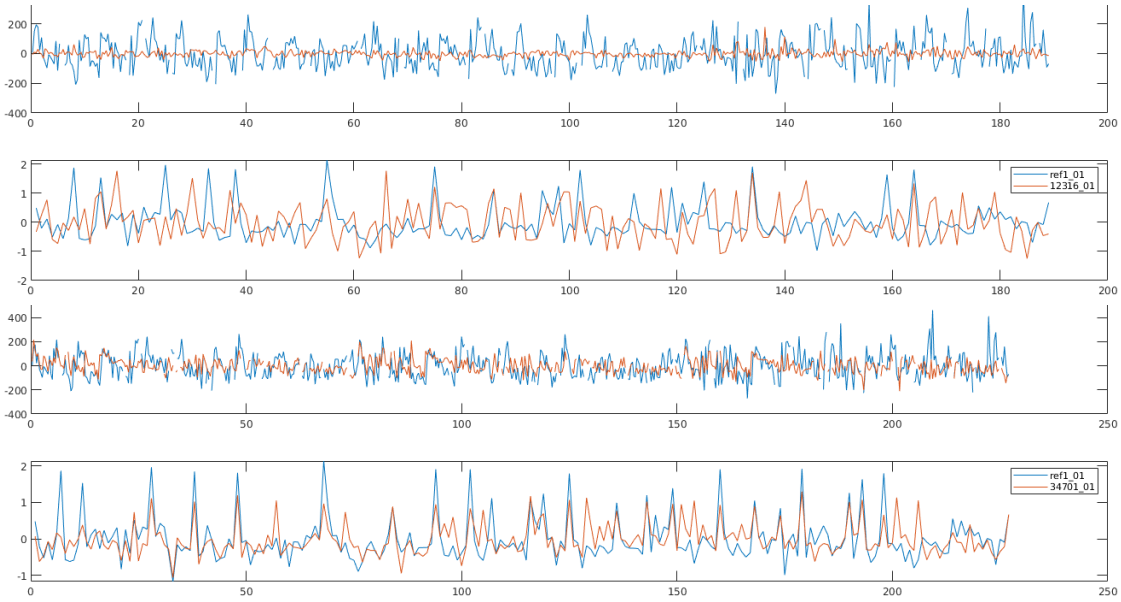


FIGURE 5.1 – Alignement de la courbe mélodique et du coefficient d'allongement des syllabes d'un lecteur à faible score à l'EMDF (en haut) et d'un autre à haut score à l'EMDF (en bas) (courbes rouges) sur un lecteur expert (courbes bleues) dans la lecture du texte 1.

utilisée par Bailly *et al.* (2010) pour l'analyse du regard : la projection des lectures de références comme des points dans un espace multidimensionnel (ici trois dimensions). La projection des lectures d'enfants dans cet espace permet alors de leur attribuer des coefficients correspondants à leur performances prosodiques. Deux performances sont prises en compte ici : les schémas mélodiques (basé sur la F0) et les schémas rythmiques (basés sur l'allongement des syllabes). Afin de s'abstraire des problèmes de registre et de vitesse, les lectures sont toutes alignées et les moyennes des réalisations sont soustraites avant d'être projetées dans les espaces de référence. La figure 5.1 illustre les alignements pour un lecteur ayant obtenu un score à l'EMDF faible, et un lecteur ayant obtenu un score élevé à l'EMDF.

Les distances entre les lectures de référence deux à deux ont été calculées pour les deux schémas prosodiques. A partir de ces distances, la mise à l'échelle multidimensionnelle permet de construire deux espaces de référence tridimensionnels.

Caractériser la prosodie des lectures d'enfants Les lectures d'enfants sont ensuite projetées dans ces espaces en deux temps. Dans un premier temps, les lectures sont alignées sur les lectures de référence, comme présenté dans la figure 5.1. La distance à chaque lecture de référence est calculée. On observe sur la figure 5.1 que le faible lecteur présente une courbe mélodique plus aplatie et une courbe d'allongement syllabique décalée par rapport au lecteur expert, engendrant une grande distance par rapport à cette référence. Le lecteur confirmé présente une courbe mélodique plus variée et des allongements syllabiques correspondant à ceux du lecteur expert, résultant en une distance plus faible à cette référence. Dans un deuxième temps, chaque distance est projetée dans l'espace de référence permettant de placer la lecture dans l'espace de reférence par rapport aux lecture de références. Cette projection permet d'obtenir alors trois coefficients pour les trois axes de chaque espace de référence : F01, F02, F03 pour l'espace mélodique et COE1, COE2 et COE3 pour l'espace rythmique. Ces six coefficients permettent de caractériser objectivement la prosodie de la lecture.

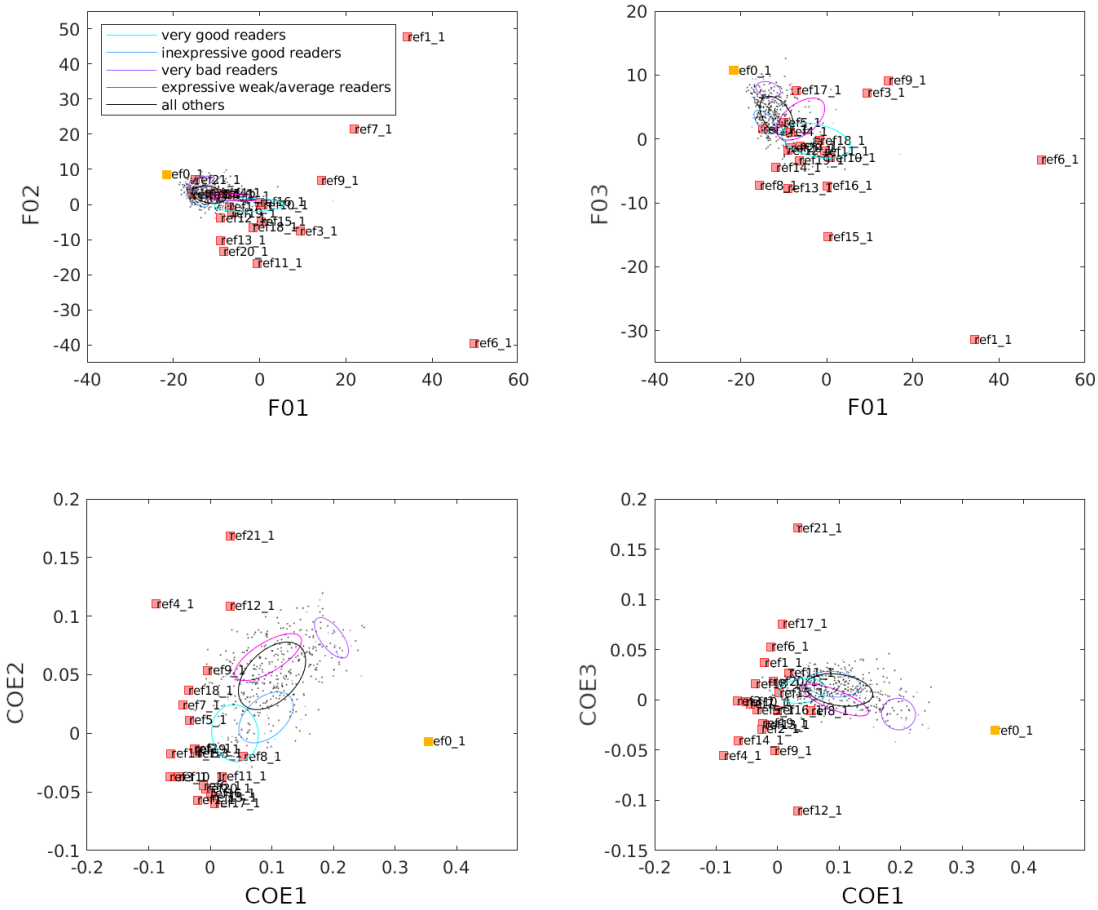


FIGURE 5.2 – Projection 2D des espaces prosodiques de référence pour le texte 1 avec les lectures de référence (carrés rouges) et celles des enfants (points) et groupes de lecteurs de profils différents Ref0 (carré orange) correspond à une lecture sans variation mélodique ou rythmique

Les figures 5.2 et 5.3 montrent la projection des lectures d'enfants dans chacun des espaces prosodiques, respectivement pour les textes 1 et 2. Dans ces représentations graphiques, nous avons ajouté des clusters de bons lecteurs, de bons lecteurs inexpressifs, de faibles lecteurs et de faibles lecteurs expressifs sur la base des évaluations subjectives effectuées par les juges. Le positionnement de ces clusters confirment l'efficacité de la méthode de projection utilisée. On constate en effet que les bons lecteurs expressifs se situent proches des références, tandis que les faibles lecteurs en sont trés éloignés, les autres catégories se situant entre les deux. On remarque également une différenciation des clusters bons lecteurs inexpressifs et faibles lecteurs expressifs, ces derniers étant plus éloignés des références. Ils sont également plus proches de la référence Ref0, référence fictive correspondant à une lecture sans variation mélodique et avec des syllabes d'égales durées

Validation de la méthode

Liens entre paramètres objectifs et évaluations subjectives Des régressions linéaires multiples ont permis de déterminer quels paramètres verbaux, prosodiques et rythmiques impactent le plus les évaluations subjectives. Le tableau 5.1 présente les régresseurs significatifs obtenus pour chaque dimension dans l'ordre d'importance. Dans une première série de régressions, seuls les paramètres verbaux ont été introduits. Les corrélation (pseudo R2 de Mac Fadden) entre scores

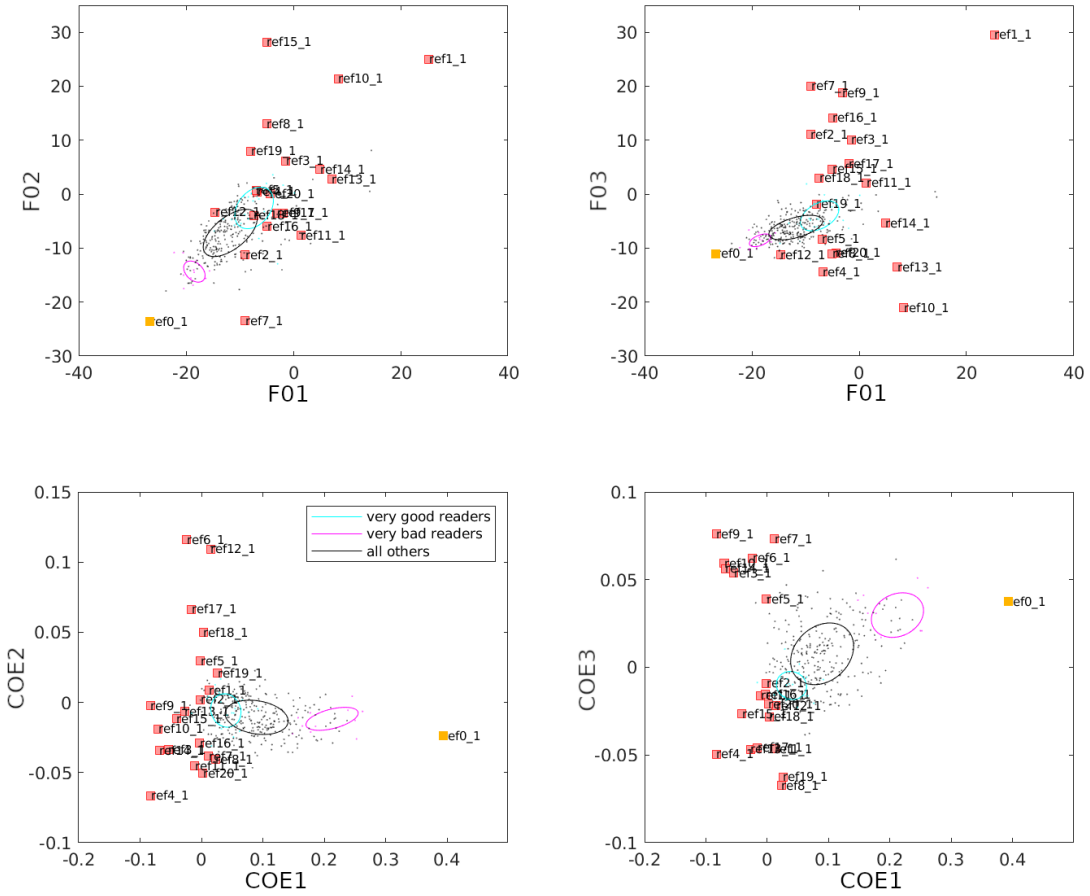


FIGURE 5.3 – Projection 2D des espaces prosodiques de référence pour le texte 1 avec les lectures de référence (carrés rouges) et celles des enfants (points) et groupes de lecteurs de profils différents
Ref0 (carré orange) correspond à une lecture sans variation mélodique ou rythmique

TABLE 5.1 – Prédicteurs principaux des scores aux évaluations subjectives et corrélation (pseudo R2 de MacFadden) pour les régressions avec paramètres verbaux uniquement et régressions avec paramètres verbaux et prosodiques. Seuls les régresseurs significatifs ($p < 0.05$)sont présentés dans l'ordre d'importance de leur contribution au modèle.

	verbal-only		**verbal+prosody**	
Text 01				
expressivité	.62	$1 + NV$	.79	$1 + F0_1 + COE_1 + NMCLM$
phrasé	.77	$1 + NMI + NV$	.83	$1 + NMCLM + COE_1$
décodage	.75	$1 + NMI + NMCLM$	.76	$1 + NMCLM + NMI$
vitesse	.82	$1 + NMCLM$	.84	$1 + NMCLM$
Text 02				
expressivité	.65	$1 + NV$	.83	$1 + COE_1 + NMCLM + F0_2 + F0_1$
phrasé	.81	$1 + NMI + NV$	.91	$1 + COE_1 + NV + COE_3 + NMI + F0_3$
décodage	.86	$1 + NMI + NV$	.88	$1 + NV + NMI + COE_1$
vitesse	.89	$1 + NV$	.91	$1 + NV + COE_1$

subjectifs et scores prédits sont alors correctes pour les dimensions d'automaticité et de phrasé et moyennes pour l'expressivité. Dans la deuxième série de régressions, l'introduction des coefficients prosodiques permet d'améliorer sensiblement ces corrélations pour les dimensions phrasé et expressivité. On constate par ailleurs que les coefficients prosodiques sont essentiellement trouvés comme régresseurs des dimensions prosodiques : coefficient rythmique pour le phrasé et l'expressivité et coefficient mélodique pour l'expressivité. Cela confirme que la méthode utilisée permet d'obtenir de bons indicateurs de la qualité prosodique de la lecture.

Finalement les corrélations entre scores subjectifs et scores prédits sont bonnes pour le texte 1 (de .76 à .84) et trés bonnes pour le texte 2 (.83 à .91) indiquant des modèles d'une bonne qualité explicative.

Dépendance au texte L'outil a été utilisé sur les textes 1 et 2 indépendamment. On peut voir dans le tableau 5.1 que les prédicteurs sont peu différents entre les deux textes. On retrouve les paramètres verbaux dans les dimensions de vitesse et de décodage et les paramètres prosodiques dans les dimensions de phrasé et d'expressivité. Il est intéressant de noter que les régressions avec uniquement les paramètres verbaux sont trés proches pour les deux textes. C'est avec l'introduction des paramètres prosodiques que les différences apparaissent. En effet, les paramètres de prosodie F0 et COE interviennent dans les dimensions de phrasé et d'expressivité mais leur importance diffère d'un texte à l'autre. Dans le texte 2, les paramètres rythmiques interviennent également dans les dimensions de décodage et de vitesse, ce qui n'est pas le cas pour le texte 1. Pour rappel, le texte 2 est un dialogue, induisant plus d'expressivité. Cette caractéristique est cohérente avec des paramètres prosodiques significatifs plus nombreux dans la régression et des corrélations plus élevées dans le texte 2. Les modèles peuvent donc différer d'un texte à l'autre mais restent cohérents dans leur structure globale combinant paramètres verbaux et prosodiques.

La figure 5.4 compare les deux textes pour les juges humains et les scores prédits. On observe pour l'expressivité des scores globalement plus élevés pour le texte 2. A l'inverse pour les autres dimensions, on constate une différenciation des tendances : les scores prédits sont plus élevés pour le texte 1 et les scores subjectifs plus élevés pour le texte 2. On peut noter que pour le phrasé cette différenciation est faible, les deux textes donnant des scores proches. De plus les scores prédits sont

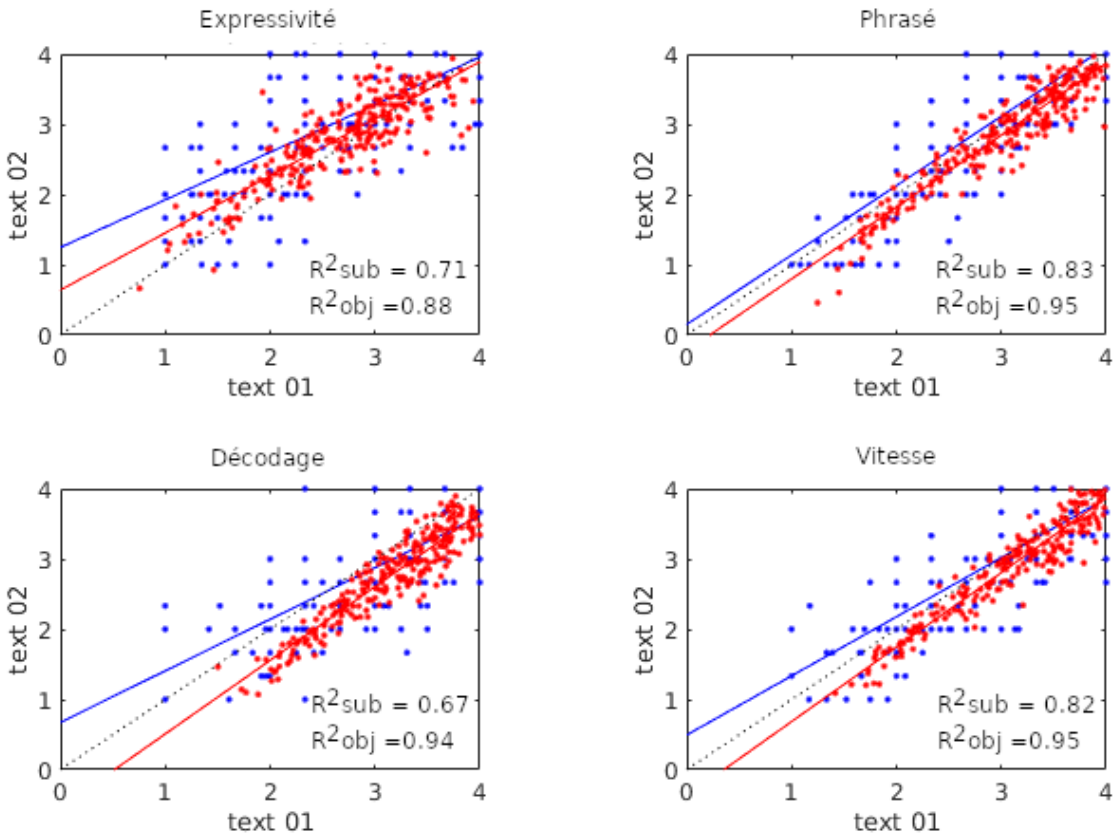


FIGURE 5.4 – Comparaison des scores moyens donnés par les juges (bleu) et scores prédits (rouge) pour les textes 1 et 2 dans les lectures du corpus transversal, pour chacune des dimensions de l'EMDF

plus consistants que les scores subjectifs. On observe en effet des R^2 trés élevés pour les scores prédits (R^2 .88). Ainsi la méthode de prédiction des scores utilisée ici peut l'être avec différents types de texte, même si chacun des textes présentent des particularités, notamment au niveau de l'expressivité.

Evaluation automatique

Le processus décrit ci-dessus a ensuite été automatisé pour évaluer directement les lectures à partir des données audio brutes. Les phases de tests ont montré que l'utilisation des paramètres prosodiques permet de réduire significativement l'erreur de prédiction.

Conclusion

Pour conclure, la méthode présentée ici utilise des paramètres verbaux et des paramètres prosodiques issus d'une comparaison avec de multiples références. Elle permet d'obtenir des scores objectifs à chaque dimensions de la fluence, ancrés dans les données acoustiques et verbales des lectures. Les scores ainsi produits sont trés corrélés aux scores subjectifs moyens donnés par de nombreux évaluateurs, tout en étant rapides et faciles à mettre en oeuvre. Ces scores sont également plus consistants entre les deux textes utilisés. Les scores générés sont des scores continus. Ces scores sont donc beaucoup plus précis que les scores discrets utilisés dans l'EMDF. Cet outil

permet donc de capter des variations plus fines d'un élève à un autre ou d'un moment à un autre. Il est important de noter que les scores dans cet outil ne sont pas bornés. Contrairement à l'EMDF, ils peuvent donc dépasser 4. Chez les bons lecteurs, c'est le cas ponctuellement, notamment sur les scores de vitesse. En effet, la vitesse est calculée à partir des paramètres verbaux (NMCLM et NV). Ainsi un lecteur beaucoup plus rapide que la moyenne de la population ayant permis d'étalonner l'outil, peut avoir des scores dépassant 4. Dans un souci de cohérence par rapport à l'échelle employée pour la construction de cet outil, nous avons choisi de ramener ces scores à 4, comme cela aurait été le cas lors d'une notation par un juge humain dans la suite de ce travail. Pour conclure, l'objectivité et la précision de cet outil lui permette d'être utilisé pour évaluer des lectures dans le cadre de la recherche. Il est particulièrement adapté aux études longitudinales. En effet, son objectivité permet de réaliser des évaluations multiples sans avoir à se soucier de la variabilité inter-évaluateur. Les scores aux évaluations répétées sont directement comparables. Sa précision permet de capter des variations fines dans les acquisitions prosodiques, uniquement mesurables avec des paramètres acoustiques.

5.1.2 Développement de la prosodie

L'outil automatique présenté dans le paragraphe précédent a été utilisé pour évaluer les lectures du corpus longitudinal. Ces scores nous donnent l'occasion d'observer le développement de chacune des dimensions de la fluence des élèves du CE1 jusqu'en CM1. Nous utilisons ici les textes 1 et 2 afin de déterminer également si le taux de croissance dépend du texte. Nous nous attendons à trouver la même trajectoire d'évolution que dans l'étude 1 : une augmentation des scores à chaque dimension avec une rupture dans le rythme d'apprentissage entre CE2 et CM1 pour les dimensions phrasé, décodage et vitesse. Les scores de la dimension expressivité devrait augmenter plus lentement.

Méthodologie - analyse statistique

Nous utilisons donc les données du corpus longitudinal évaluées par l'outil présenté ci-dessus. Cette étude concerne les 49 enfants qui ont été enregistrés lors des trois temps du suivi longitudinal. Les scores prédits moyens pour chaque année et chaque texte sont présentés dans le tableau 5.2. L'outil de prédiction automatique a donné six scores supérieurs à 4 en CM1 (compris entre 4.01 et 4.15 à la dimension vitesse). Pour rester cohérent avec l'échelle utilisée dont les scores sont compris entre 1 et 4, nous avons choisi de ramener ces trois scores à 4, score maximal qui serait obtenu avec des évaluateurs humains.

Toutes les analyses statistiques ont été réalisées avec le logiciel R (Team, 2019). L'effet du niveau a été testé en utilisant des modèles linéaires multiniveaux de croissance, en utilisant le sujet comme paramètre aléatoire de niveau 2 et le parcours comme paramètre aléatoire de niveau 3, avec la fonction lmer du package lme4 (Bates *et al.*, 2015). Les participants sont répartis sur plusieurs classes chaque année (quatre classes en CE1, cinq en CE2 et six en CM1). Cette répartition évolue chaque année. Les sujets suivent donc des parcours différents pendant ces trois années. Au total les participants se répartissent sur 12 parcours du CE1 au CE2 (année 2) et sur 25 parcours du CE1 au CM1 (année 3). Pour chaque modèle, l'effet aléatoire de niveau 3 est significatif sur le modèle vide. Il est donc conservé pour la suite des analyses. Quand l'effet du niveau est significatif, une analyse post-hoc a été menée pour déterminer les différences entre les niveaux deux à deux. Nous avons utilisé le test HSD de Tukey. Pour cette analyse nous rapportons la p-value la plus élevée parmi les tests deux à deux. Nous avons également testé l'effet du texte et son interaction avec le niveau.

TABLE 5.2 – Scores prédits à l'EMDF pour le texte 1 par l'outil automatique pour chaque année du corpus longitudinal

	Texte 1			Texte 2		
année	1	2	3	1	2	3
niveau	CE1	CE2	CM1	CE1	CE2	CM1
Expressivité	2.16(0.53) 1.34 : 3.59	2.35(0.53) 1.12 : 3.54	2.57(0.52) 1.04 : 3.75	2.22(0.37) 1.13 : 3.6	2.55(0.42) 1.62 : 3.57	2.82(0.46) 1.64 : 3.97
Phrasé	2.35(0.49) 1.45 : 3.51	2.71(0.50) 1.47 : 3.58	2.99(0.50) 1.35 : 3.85	2.21(0.44) 1 : 3.37	2.71(0.43) 1.53 : 3.78	3.03(0.42) 1.75 : 3.94
Décodage	2.62(0.43) 1.5 : 3.71	2.92(0.39) 1.92 : 3.62	3.10(0.40) 1.83 : 3.85	2.19(0.33) 1.38 : 2.87	2.60(0.35) 3.41 : 2.92	2.91(0.37) 2.09 : 3.94
Vitesse	2.53(0.49) 1.71 : 3.92	2.98(0.49) 1.62 : 3.84	3.24(0.48) 1.7 : 4	2.32(0.41) 1.34 : 3.6	2.83(0.44) 1.61 : 3.98	3.14(0.40) 2.03 : 4

TABLE 5.3 – Modèle de croissance du CE1 au CM1 à la dimension expressivité

	Modèle 1	**Modèle 2**	**Modèle 3**	**Modèle 4**
Effets fixes				
Constante	2.33(0.08)***	1.86(0.09)***	1.86(0.09)***	1.89(0.09)***
niveau		0.25(0.02)***	0.25(0.02)***	0.22(0.02)***
Texte2				-0.04(0.08)
niveau*texte				0.08(0.04)*
Effets aléatoires				
Niveau 3 (parcours)	0.06(0.25)	0.07(0.26)	0.09(0.25)	0.06(0.25)
Niveau 2 (élèves)				
- Variance des constantes	0.15(0.39)	0.17(0.41)	0.16(0.40)	0.16(0.40)
- Variance des pentes			0.01(0.11)	
Qualité de l'ajustement				
AIC	364	232	233	226

Résultats

La figure 5.5 présente l'évolution des scores avec le niveau scolaire pour chaque dimension et chaque texte. On peut remarquer sur cette figure le décalage de scores entre les deux textes et également la différence de taux de croissance. Ces observations sont confirmées par les modèles de croissance présentés ci-dessous.

Expressivité Le modèle de croissance présenté dans le tableau 5.3 montre un effet fixe positif du niveau sur le score à la dimension expressivité. L'ajout du niveau comme paramètre aléatoire de niveau 2 n'améliore pas notre modèle. On observe donc un effet du niveau mais pas de variance des pentes entre les élèves. Dans le modèle 4, l'interaction entre niveau et texte améliore notre modèle et montre un taux de croissance légèrement plus élevé pour le texte 2 (0.08 point en plus par an que le texte 1), mais pas d'effet significatif du texte sur les scores. Comme on peut le constater sur la figure 5.5, la croissance des scores à la dimension expressivité semble assez linéaire du CE1 au CM1. L'analyse post-hoc de Tukey confirme la croissance progressive de l'expressivité avec des scores significativement différents pour chaque année ($p < .001$). Les scores à la dimension expressivité augmente en moyenne de 0.25 points par an.

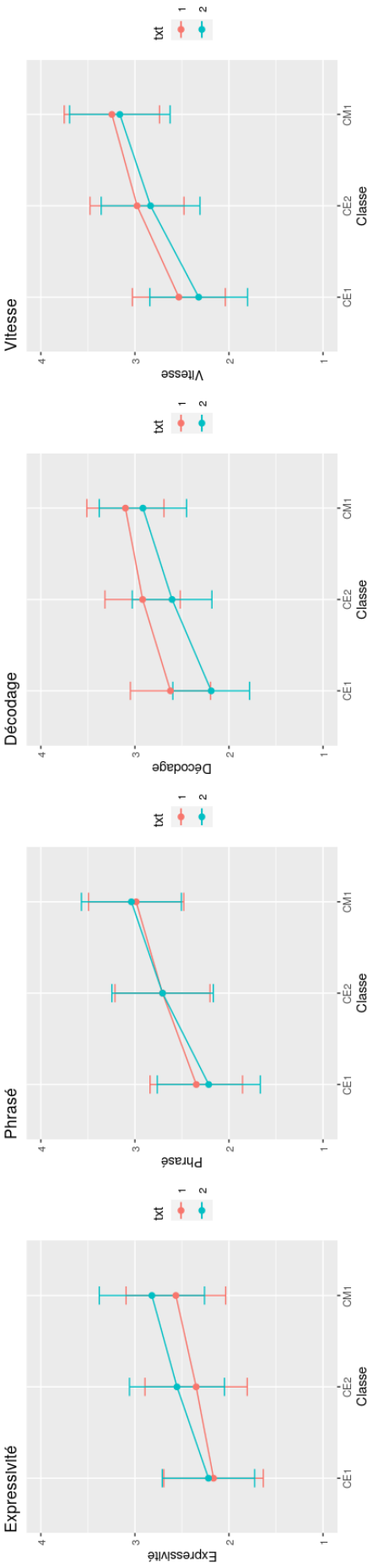


FIGURE 5.5 – Croissance des scores, évalués par outil automatique, aux quatres dimensions de l'EMDF pour les élèves suivis du CE1 au CM1

Table 5.4 – Modèle de croissance du CE1 au CM1 à la dimension phrasé

	Modèle 1	**Modèle 2**	**Modèle 3**	**Modèle 4**
Effets fixes				
Constante	2.61(0.08)***	1.86(0.09)***	1.84(0.10)***	1.94(0.11)***
niveau		0.37(0.02)***	0.37(0.02)***	0.34(0.02)***
Texte2				-0.23(0.06)**
niveau*texte				0.07(0.03)*
Effets aléatoires				
Niveau 3 (parcours)	0.09(0.29)	0.09(0.30)	0.14(0.38)	0.15(0.39)
Niveau 2 (élèves)				
- Variance des constantes	0.12(0.35)	0.15(0.38)	0.14(0.38)	0.15(0.39)
- Variance des pentes			0.007(0.08)	0.009(0.09)
Qualité de l'ajustement				
AIC	432	146	144	138

Table 5.5 – Modèle de croissance du CE1 au CM1 à la dimension décodage

	Modèle 1	**Modèle 2**	**Modèle 3**	**Modèle 4**
Effets fixes				
Constante	2.70(0.06)***	2.09(0.08)***	2.33(0.10)***	
niveau		0.30(0.02)***	0.31(0.02)***	0.26(0.02)***
Texte2				-0.57(0.06)***
niveau*texte				0.10(0.03)*
Effets aléatoires				
Niveau 3 (parcours)	0.04(0.20)	0.05(0.22)	0.10(0.31)	0.06(0.24)
Niveau 2 (élèves)				
- Variance des constantes	0.07(0.27)	0.09(0.30)	0.05(0.22)	0.10(0.32)
- Variance des pentes			0.001(0.04)	
Qualité de l'ajustement				
AIC	393	230	234	55

Phrasé Le modèle de croissance présenté dans le tableau 5.4 montre un effet fixe positif du niveau sur le score à la dimension phrasé. L'ajout du niveau comme paramètre aléatoire de niveau 2 améliore notre modèle. On observe donc un effet du niveau et une variance des pentes entre élèves. Le rythme de progression est donc différent entre les élèves. Dans le modèle 4, l'interaction entre niveau et texte améliore notre modèle. On constate à la fois un effet du texte sur les scores, significativement plus faibles pour le texte 2 (-0.23 points), et sur le taux de croissance légèrement plus élevé pour le texte 2 (0.07 points en plus par an). Comme on peut le constater sur la figure 5.5.B la croissance des scores à la dimension phrasé semble assez linéaire du CE1 au CM1. L'analyse post-hoc de Tukey confirme la croissance progressive de l'expressivité avec des scores significativement différents pour chaque année ($p < .001$). Les scores à la dimension expressivité augmentent en moyenne de 0.37 points par an.

Décodage Le modèle de croissance présenté dans le tableau 5.5 montre un effet fixe positif du niveau sur le score à la dimension décodage. L'ajout du niveau comme paramètre aléatoire de niveau 2 n'améliore pas notre modèle. On observe donc une effet du niveau mais pas de différence

Table 5.6 – Modèle de croissance du CE1 au CM1 à la dimension vitesse

	Modèle 1	**Modèle 2**	**Modèle 3**	**Modèle 4**
Effets fixes				
Constante	2.80(0.08)***	2.02(0.09)***	2.03(0.10)***	2.15(0.11)***
niveau		0.39(0.02)***	0.39(0.02)***	0.37(0.02)***
Texte2				-0.3(0.005)***
niveau*texte				0.04(0.02)*
Effets aléatoires				
Niveau 3 (parcours)	0.07(0.26)	0.08(0.27)	0.15(0.38)	0.16(0.39)
Niveau 2 (élèves)				
- Variance des constantes	0.13(0.36)	0.15(0.39)	0.12(0.34)	0.14(0.38)
- Variance des pentes			0.003(0.06)	0.008(0.08)
Qualité de l'ajustement				
AIC	449	151	148	71

entre élèves dans les rythmes d'acquisition. Dans le modèle 4, l'ajout de l'interaction du niveau avec le texte améliore grandement notre modèle. Les scores du textes 2 sont significativement inférieurs à ceux du texte 1 (-0,57 points) et son taux de croissance significativement supérieur (0.10 points de plus par an). Comme on peut le constater sur la figure 5.5.C la croissance des scores à la dimension décodage semble changer du CE2 au CM1. Les scores, compris entre 1.83 et 3.85, pour le texte 1, ne plafonnnent pas en CM1. L'analyse post-hoc de Tukey confirme la croissance progressive du décodage avec des scores significativement différents pour chaque année ($p<.001$). Les scores à la dimension décodage augmentent en moyenne de 0.31 points par an.

Vitesse Le modèle de croissance présenté dans le tableau 5.6 montre un effet fixe positif du niveau sur le score à la dimension vitesse. L'ajout du niveau comme paramètre aléatoire de niveau 2 améliore notre modèle. On observe donc un effet fixe du niveau et une variance des pentes au niveau inter-élèves. Tous les élèves ne progressent donc pas au même rythme. Dans le modèle 4, l'ajout de l'interaction entre niveau et texte améliore le modèle. Le score à la dimension vitesse est significativement plus faible pour le texte 2 (-0.30 points) et le taux de croissance légèment plus élevé. La différence reste cependant faible (+ 0.04 points par an). Comme on peut le constater sur la figure 5.5.D la croissance des scores à la dimension vitesse semble ralentir du CE2 au CM1. Les scores, compris entre 1.7 et 4 plafonnnent en CM1. L'analyse post-hoc HSD de Tukey confirme la croissance progressive de la vitesse avec des scores significativement différents pour chaque année ($p < .001$). Les scores à la dimension vitesse augmentent en moyenne de 0.39 points par an.

Discussion

Ces modèles de croissance des différentes dimensions de la fluence viennent confirmer les résultats des modèles de développement obtenus sur les données du corpus transversal. On peut constater que les scores à toutes les dimensions augmentent du CE1 au CM1. Les scores d'origine varient d'une dimension à l'autre, avec des scores plus élevés dans les dimensions d'automaticité, que dans les dimensions prosodiques. Le rythme d'acquisition moyen dépend également des dimensions. Les scores de vitesse et de phrasé augmentent plus rapidement que ceux de décodage et d'expressivité. Les raisons en sont probablement différentes. Le score de décodage progresse plus

lentement car les enfants de CE1 ont déjà un bon niveau de décodage. L'expressivité est encore en cours d'acquisition et est une compétence plus difficile et lente à acquérir. Le phrasé progresse plus rapidement. Il est possible que la fenêtre de temps à laquelle nous nous sommes intéressés ici cible une période de phrasé en développement, décodage acquis et vitesse en augmentation. On remarque, comme dans les autres chapitres, l'exception de la dimension expressivité qui semble plus tardive et longue à acquérir que les autres dimensions.

Cette étude nous apporte également d'autres informations sur les rythmes d'acquisition de ces compétences. Pour l'expressivité et le décodage, le rythme d'acquisition est le même pour tous les enfants suivis. On observe des profils de développement différents suivant les enfants pour les dimensions de phrasé et de vitesse. Il serait donc intéressant ici de tester différentes compétences associées, comme le vocabulaire ou la compréhension par exemple, qui pourraient expliquer ces différences inter-individuelles. Il peut paraitre étonnant que ces différences de rythme de progression n'apparaisse que dans certaines dimensions. Pour la vitesse, cela pourrait s'expliquer par les scores élevés en CM1. En effet, pour les bons lecteurs, l'automaticité est acquise dès la fin du CE1 et ces textes ne posent normalement que peu de difficulté de décodage. Ces lecteurs ont donc des scores élevés dés le CE1 et n'ont donc que peu de marge de progression. A l'inverse, les lecteurs plus faibles ont une marge de progression plus forte entre le CE1 et le CM1, et donc un rythme de progression plus élevé. Pour la prosodie, les scores de départ sont plus faibles, particulièrement pour l'expressivité, y compris pour les bons lecteurs. La progression se fait donc pour tous à un même rythme. Ainsi il serait intéressant de confirmer cette hypothèse en séparant les sujets en groupe suivant leur score initial en automaticité et en vérifiant l'interaction du groupe avec le rythme de progression des scores.

La croissance est linéaire pour les dimensions prosodiques, avec des scores qui n'atteignent pas le plafond. Le choix du modèle linéaire semble alors judicieux dans notre cas. Cependant au vu des résultats présentés dans le chapitre 2.1 et des scores obtenus par les CM1 dans les dimensions d'automaticité, il est probable que les scores de décodage et de vitesse tendent vers un plafond en CM1. Le modèle linéaire qui semble encore convenir sur les trois points de mesure du CE1 au CM1 n'est peut-être pas le plus adapté. En effet, avec l'effet plafond possible à partir du CM1, il est probable qu'un modèle non linéaire soit plus adapté pour modéliser l'effet plafond, de type exponentielle inverse. L'utilisation d'une fenêtre de développement plus longue (CP au CM2) permettrait d'avoir une meilleure idée du type de modèle à utiliser.

La modélisation en utilisant les deux textes permet d'observer une différence dans les rythmes de progression. Le texte 2 présente un taux de croissance plus élevé dans toutes les dimensions. Cette différence est particulièrement forte pour les dimensions d'expressivité et de décodage. Le texte 2 étant un dialogue, il induit plus d'expressivité. Avec le temps et l'acquisition du décodage, les enfants sont de plus en plus à l'aise pour "mettre en voix" le dialogue. Il est également possible que l'exercice en CM1 soit plus connu et pratiqué en classe qu'en CE1. Cette différence dans le taux de croissance est étonnante pour le décodage déjà en place chez la majorité des enfants en CE1. Il est possible que le format dialogue soit plus difficile à appréhender et donne plus d'hésitations en CE1 car il comprend des phrases courtes et des interjections sur lesquelles les enfants butent et se reprennent, entrainant des ruptures de rythme et des répétitions et donc des scores de décodage plus faibles. Avec la familiarisation à ce type de texte, ils hésitent moins et les scores de décodage augmentent plus fortement que pour le texte 1, narratif, et plus classique.

Il serait intéressant de prolonger cette étude longitudinale pour avoir plus d'années de mesure et mieux cerner le profil d'acquisition de chacune des dimensions et peut-être les liens entre ces

profils. En considérant que l'automaticité est une pré-requis à la prosodie, on s'attendrait à une acquisition rapide du décodage et de la vitesse en CP/CE1 suivi d'un ralentissement une fois ces compétences maitrisées. A l'inverse on peut s'attendre à une acquisition plus lente pour le phrasé et l'expressivité, qui s'accélère une fois que les acquisitions des dimensions d'automaticité ralentissent. Une étude sur une durée plus longue, nous permettrait également d'avoir plus d'informations sur les différents profils de développement et leurs origines possibles.

Résultats majeurs **Étude 6**

L'outil d'analyse automatique de la fluence à partir des paramètres verbaux et prosodiques est un outil valide et d'une bonne précision qui permet de s'affranchir de la variabilité liée aux juges.

Du CE1 au CM1...

L'acquisition de l'expressivité est plus lente et tardive que celles des autres dimensions de la fluence mesurée avec l'outil automatique.

Le rythme d'acquisition diffère d'un enfant à l'autre pour le phrasé et la vitesse.

Les scores à toutes les dimensions augmentent plus rapidement pour le dialogue que pour le texte narratif, particulièrement pour l'expressivité et le décodage.

5.2 Étude 7 : Etude longitudinale du développement des dimensions prosodiques de la lecture et compréhension écrite

Dans les chapitres 3 et 4, nous avons caractérisé le développement de la prosodie en lecture d'un point de vue transversal et le lien entre prosodie et compréhension au sein d'un même niveau scolaire. Dans l'étude précédente, nous avons pu observer le développement de chacune des dimensions de la fluence. Dans cette dernière étude, nous nous intéressons aux liens de causalité entre les quatre dimensions de la fluence et la compréhension écrite. Comme mentionné dans le chapitre 1, il existe peu d'études ayant exploré la causalité des liens dans le développement des dimensions prosodiques et de la compréhension. L'étude de Miller et Scwanenflugel (2008) s'intéressant au développement longitudinal de la prosodie en lecture utilise un nombre de paramètres acoustiques limités (nombre de pauses et variations de F0) et se limite aux toutes premières année de la lecture (CP au CE2). Dans cette étude, les auteures ont pu montrer que les variations de F0 en CP prédisaient le nombre de pauses en CE1 et la compréhension en CE2. Les autres études longitudinales s'intéressant à la compréhension utilisent toutes une échelle d'évaluation subjective de la fluence, la MDFS (Rasinski, 2004) ou ses traductions. Elles s'intéressent à des langues et niveaux différents et leurs résultats divergent. L'hypothèse globale qui se dégage de la synthèse de ces études (cf. chapitre 1) est un lien de causalité bidirectionnel variant au cours de l'apprentissage. Il serait intéressant de confirmer et approfondir les résultats de ces études. En effet, elles ne prennent en compte que le score global à l'échelle subjective, regroupant indifféremment toutes les dimensions de prosodie et d'automaticité en un seul score. Or on peut supposer que toutes ces dimensions n'ont pas les mêmes liens avec la compréhension.

Dans cette dernière étude, nous souhaitons donc mettre en évidence les liens de causalité dans le développement du décodage, de la vitesse, du phrasé, de l'expressivité et de la compréhension écrite. Nous explorons ici les relations longitudinales à la fois entre les différentes dimensions de la fluence et leur lien avec la compréhension écrite. Nous nous attendons à trouver des liens bidirectionnels qui varient au fil du développement.

5.2.1 Méthodologie

Données

Pour cette étude, nous utilisons les données du corpus longitudinal. Seules les données du texte 1 sont utilisées ici. Les participants sont décrits dans le tableau 2.2. Les données utilisées pour la modélisation sont celles des 50 participants en CE1 et CE2 et des 53 participants en CE2 et CM1 qui ont été enregistrés et ont participé aux tests écrits.

Prosodie Nous utilisons pour cette étude les résultats de l'évaluation automatique des lectures présentée dans le paragraphe précédent. Nous prenons en compte les scores prédits à chaque dimension de l'échelle : expressivité, phrasé, décodage et vitesse. Ces scores prennent des valeurs continues comprises entre 1 et 4. L'outil de prédiction automatique a cependant donné six scores supérieurs à 4 en CM1 (compris entre 4.01 et 4.15 à la dimension vitesse). Pour rester cohérent avec l'échelle utilisée dont les scores sont compris entre 1 et 4, nous avons choisi de ramener ces scores à 4, score maximal qui serait obtenu avec des évaluateurs subjectifs. Les scores moyens pour chaque année de l'étude sont présentés dans le tableau 5.2.

Table 5.7 – Scores moyens aux tests de compréhension et scores moyens aus dimensions de la fluence prédits par l'outil automatique pour chaque année du corpus longitudinal

Année	1	2	3
Niveau	CE1	CE2	CM1
Compréhension globale	.71(.23)	.68(.23)	.66(.20)
Compréhension locale	-	.92(.09)	.75(.18)
Compréhension en-ligne	.16(.12)	.15(.10)	.15(.09)
Expressivité	2.16(0.53)	2.35(0.53)	2.57(0.52)
Phrasé	2.35(0.49)	2.71(0.50)	2.99(0.50)
Décodage	2.62(0.43)	2.92(0.39)	3.10(0.40)
Vitesse	2.53(0.49)	2.98(0.49)	3.24(0.48)

Compréhension écrite Nous utilisons les trois scores de compréhension présentés dans le chapitre 2 : compréhension globale, compréhension locale et compréhension en-ligne. Les résutats moyens obtenus à ces tests sont présentés dans le tableau 5.7.

Variables contrôlées Nous contrôlons dans ces analyses les scores de vocabulaire, raisonnement non-verbal, compréhension orale et NMCLM.

Analyses statistiques

Pour construire notre modèle longitudinal de développement de la prosodie en lecture et de la compréhension écrite, nous avons choisi d'utiliser un modèle linéaire général multiniveaux. Dans les régressions, nous prenons en compte l'effet aléatoire de ces parcours à un niveau 2. En effet, les participants sont répartis sur plusieurs classes chaque année (quatre classes en CE1, cinq en CE2 et six en CM1). Cette répartition évolue chaque année. Les sujets suivent donc des parcours différents pendant ces trois années. Au total les participants se répartissent sur 12 parcours du CE1 au CE2 (année 2) et sur 25 parcours du CE1 au CM1 (année 3).

Nous avons réalisé deux séries de régressions : de l'année 1 vers l'année 2 et de l'année 2 vers l'année 3. Cela permet de simplifier les modèles et de maximiser le nombre de participants puisque que nous pouvons ainsi prendre en compte les participants présents sur deux années uniquement. Les régressions ont été réalisées sur les variables suivantes pour les années n (2 et 3) :

- Compréhension écrite globale
- Compréhension écrite locale
- Compréhension écrite en-ligne
- Expressivité
- Phrasé
- Décodage
- Vitesse

Les régresseurs étaient les mêmes variables pour les années n-1 (1 et 2), additionnées des variables contrôles (NMCLM, vocabulaire, raisonnement non verbal et compréhension orale).

Tous les scores étant compris entre 0 et 1, nous avons réalisé des régressions bêta en utilisant la fonction glmmadmb du package *glmmADMB* (Fournier *et al.*, 2012). Nous avons réalisé des régressions multiples, en comparant par une anova les modèles successifs et en conservant les paramètres les plus significatifs à chaque étape, avant de calculer les coefficients de régression et les coefficients de détermination pour chaque variable indépendante significative du modèle. Nous rapportons la part de variance propre des variables d'intérêt, c'est à à dire leur part de variance, une fois retirée la variance expliquée par les variables contrôles.

5.2.2 Résultats

Du CE1 au CE2

Prédiction des dimensions prosodiques de la lecture Le tableau 5.8 présente les résultats des régressions des dimensions d'expressivité et de phrasé.

Pour l'expressivité, une fois le vocabulaire controlé, on observe un effet significatif du décodage et de l'expressivité en CE1. Il n'y a pas d'effet significatif de la compréhension en CE1. Le décodage en CE1 explique une part propre de variance importante (32%) du score d'expressivité en CE2. Le score d'expressivité en CE1 explique également une part propre de variance, quoique plus faible (2.2%).

Pour le phrasé, une fois le vocabulaire controlé, on observe un effet significatif des scores de phrasé et de vitesse en CE1. Le score de phrasé de CE1 explique une part propre importante (34.1%) de la variance du score de phrasé en CE2. Le score de vitesse en CE1 explique également une part propre de variance, plus faible (1.3%).

Les résultats des régressions des scores de décodage et vitesse sont présentés dans le tableau 5.9.

Pour le décodage, une fois controlés le vocabulaire et le NMCLM, on observe un effet significatif du score de vitesse en CE1. Il n'y a aucun effet significatif des scores de compréhension. Le score de vitesse de CE1 apporte une part de variance propre du score de vitesse en CE2 assez faible(2.8%).

Pour la vitesse, seul le score de vitesse en CE1 a un effet significatif. Celui-ci explique une part de variance propre trés importante du score de vitesse en CE2, de 53.7%. Là encore les scores de compréhension n'ont pas d'effet significatif.

Prédiction de la compréhension Les résultats des régressions sur les scores de compréhension globale et locale sont présentés dans le tableau 5.10. Les régressions sur les scores de compréhension en-ligne n'ont abouti à aucun effet significatif des régresseurs testés. Pour la compréhension écrite globale, une fois le vocabulaire, le raisonnement non-verbal et la compréhension orale controlés, on n'observe aucun effet significatif des scores aux différentes dimensions prosodiques et des scores de compréhension en CE1. Finalement, pour la compréhension écrite locale, une fois le vocabulaire controlé, on observe un effet de l'expressivité en CE1. L'expressivité en CE1 apporte donc une part de variance propre, de 11.3%, au score de compréhension de CE2. On retrouve donc bien le lien de la prosodie vers la compréhension dans le développement des compétences en lecture du CE1 vers le CE2.

Pour conclure, dans les régressions sur les scores de CE2, seul apparait le lien prosodie vers compréhension, et plus précisément le score à la dimension expressivité prédit le score de compréhension locale. Les différentes dimensions de l'échelle de fluence étant essentiellement prédites par les les scores à ces mêmes dimensions en CE1.

TABLE 5.8 – Modèles multiniveaux pour les scores d'expressivité et de phrasé de l'année 2

	Expressivité			**Phrasé**		
	Modèle 1	**Modèle 2**	**Modèle 3**	**Modèle 1**	**Modèle 2**	**Modèle 3**
Effets fixes						
Constante	-0.83(0.21)***	-2.5(0.3)***	-2.26(0.28)***	-0.43(0.24)	-1.63(0.16)***	-1.81(0.17)***
Vocabulaire	0.08(0.01)***	0.04(0.01)***	0.036(0.011)**	0.081(0.016)***	0.026(0.01)**	0.028(0.009)**
Expressivité			1.20(0.57)*			
Phrasé					3.48(0.34)***	1.82(0.02)*
Décodage		3.41(0.47)***	2.19(0.72)**			
Vitesse						1.78(0.79)*
Effets aléatoires						
Niveau 2 (classe)	2.1e-9(4.6e-5)	6.1e-9(7.8e-5)	2.5e-8(1.6e4)	8.9e-3(0.095)	2.4e-9(4.9e-5)	3.9e-9(6.2e-5)
Qualité de l'ajustement						
AIC	-82.6	-117.4	-119.6	-93.2	-149.3	-152.1
adjR^2	.408	.728	.75	.502	.841	.854
ΔadjR^2	-	.32	.022	-	.341	.013

TABLE 5.9 – Modèles multiniveaux pour les scores de décodage et de vitesse de l'année 2

	Décodage		**Vitesse**	
	Modèle 1	**Modèle 2**	**Modèle 1**	**Modèle 2**
Effets fixes				
Constante	-0.36(0.14)**	-0.75(0.23)***	1.09(0.13)***	-1.98(0.23)***
Vocabulaire	0.029(0.009)***	0.02(0.01)*		
NMCLM	7.9e-3(0.9e-4)***	0.005(0.001)**		
Vitesse		1.48(0.68)*		5.02(0.37)***
Effets aléatoires				
Niveau 2 (classe)	1e-8(1e-4)	2.7e-9(5.2e-5)	0.07(0.25)	2.37e-8(1.5e-4)
Qualité de l'ajustement				
AIC	-152.2	-154.8	-73.7	-146.9
adjR^2	.739	.767	.273	.81
ΔadjR^2	-	.028		.537

Table 5.10 – Modèles multiniveaux pour les scores de compréhension globale et locale de l'année 2

	CE globale	**CE locale**	
	Modèle 1	**Modèle 1**	**Modèle 2**
Effets fixes			
Constante	-2.25(0.44)***	0.94(0.43)*	0.10(0.68)
Vocabulaire	0.10(0.03)**	0.12(0.03)***	0.091(0.035)**
RNV	0.027(0.011)*		
CO	0.13(0.06)*		
Expressivité			2.32(1.27)*
Effets aléatoires			
Niveau 2 (classe)	2.25e-8(1.5e-4)	9.9e-8(3.1e-4)	0.043(0.207)
Qualité de l'ajustement			
AIC	-49.5	-159.8	-161.6
adjR^2	.58	.285	.398
ΔadjR^2			.113

Du CE2 au CM1

Prédiction des dimensions prosodiques de la lecture Le tableau 5.11 présente les résultats des régressions des dimensions d'expressivité et de phrasé.

Pour l'expressivité, une fois le NMCLM controlé, on observe uniquement un effet significatif de l'expressivité en CE2. L'expressivité en CE2 explique une part propre de variance importante (33.1%) du score d'expressivité en CM1.

Pour le phrasé, une fois le NMCLM controlé, on observe un effet significatif du score de phrasé en CE2. Ce score explique une part propre de la variance du score de phrasé en CM1 de 7.5%.

Le tableau 5.12 présente les résultats des régressions des dimensions d'expressivité et de phrasé.

Pour le décodage, une fois controlé le NMCLM, on observe un effet significatif du score de phrasé en CE2. Ce score apporte une part de variance propre du score de vitesse en CE2 de 10%.

Pour la vitesse, une fois controlé le NMCLM, on observe un effet significatif du score de vitesse en CE2. Celui-ci explique une part de variance propre importante du score de vitesse en CM1, de 10.1%. Là encore les scores de compréhension n'ont pas d'effet significatif.

Prédiction de la compréhension Les résultats des régressions sur les scores de compréhension globale et locale sont présentés dans le tableau 5.13.

Les régressions sur les scores de compréhension en-ligne n'ont abouti à aucun effet significatif des régresseurs testés, un fois controlé la compréhension orale. Pour la CE globale, une fois le NMCLM controlé, on observe uniquement un effet significatif du score de CE globale en CE2. Ce score explique une part de variance propre de 33.1% du score de CE globale en CM1.

Pour la compréhension écrite locale, on observe un effet de la compréhension écrite globale et du score à la dimension phrasé en CE2. La CE globale en CE2 apporte donc une part de variance

TABLE 5.11 – Modèles multiniveaux pour les scores d'expressivité et de phrasé de l'année 3

	Expressivité		**Phrasé**	
	Modèle 1	**Modèle 2**	**Modèle 1**	**Modèle 2**
Effets fixes				
Constante	0.63(0.09)***	-1.30(0.23)	-0.48(0.20)*	-1.60(0.26)***
NMCLM			9.4e-3(1.1e-3)***	0.0038(0.0014)**
Expressivité		3.27(0.38)***		
Phrasé				3.08(0.59)***
Effets aléatoires				
Niveau 2 (classe)	0.029(0.17)	3.1e-9(5.5e-5)	0.0.025(0.16)	0.01(0.10)
Qualité de l'ajustement				
AIC	-64	-107.6	-123.8	-144.4
adjR^2	.26	.591	.725	.80
ΔadjR^2	-	.331	-	.075

TABLE 5.12 – Modèles multiniveaux pour les scores de décodage et de vitesse de l'année 3

	Décodage		**Vitesse**	
	Modèle 1	**Modèle 2**	**Modèle 1**	**Modèle 2**
Effets fixes				
Constante	-0.26(0.17)	-1.07(0.23)***	-0.71(0.24)**	-2.03(0.37)***
NMCLM	.0088(0.0009)***	0.0043(0.0013)***	0.013(0.001)***	0.006(0.002)**
Phrasé		2.35(0.52)***		
Vitesse				3.45(0.81)***
Effets aléatoires				
Niveau 2 (classe)	7.9e-8(2.8e-4)	6.2e-8(2.5e-4)	2.8e-4(0.017)	9.5e-8(3.1e-4)
Qualité de l'ajustement				
AIC	-142.5	-157.7	-130.7	-144.7
adjR^2	.658	.758	.69	.791
ΔadjR^2	-	.10	-	.101

TABLE 5.13 – Modèles multiniveaux pour les scores de compréhension globale et locale de l'année 3

	CE globale		**CE locale**		
	Modèle 1	**Modèle 2**	**Modèle 1**	**Modèle 2**	**Modèle 3**
Effets fixes					
Constante	-1.1(0.4)**	-1.81(0.31)***	1.11(0.15)***	-0.68(0.29)*	-1.77(0.51)***
NMCLM	9.7e-3(2e-3)***	0.0042(0.0018)*			
CE globale		2.51(0.43)***		2.68(0.39)***	2.08(0.44)***
Phrasé					2.20(0.86)*
Effets aléatoires					
Niveau 2 (classe)	9.8e-9(9.9e-5)	9e-8(3e-4)	0.09(0.31)	0.09(0.29)	0.08(0.28)
Qualité de l'ajustement					
AIC	-43.7	-69.3	-47.5	-79.6	-83.8
adjR^2	.378	.709	.29	.638	.677
ΔadjR^2	-	.331	-	.348	.039

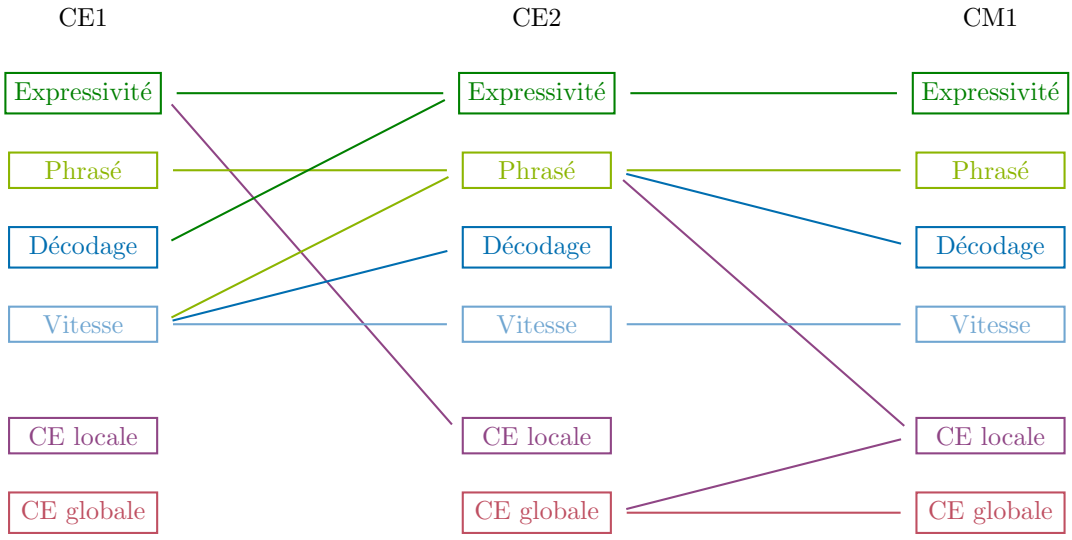


FIGURE 5.6 – Schéma des liens longitudinaux entre compréhension et dimensions prosodiques de la lecture orale du CE1 au CM1 - Chaque couleur correspond à un modèle de l'année n-1 vers n.

propre, de 34.8%, au score de CE locale de CM1. Le phrasé apporte également une part propre de variance, plus faible, de 3.9%. On retrouve donc bien le lien de la prosodie vers la compréhension dans le développement des compétences en lecture du CE2 vers le CM1.

Pour conclure, on retrouve le lien prosodie vers compréhension écrite du CE2 vers le CM1, plus précisément le phrasé en CE2 explique la compréhension locale en CM1. On ne retrouve par contre aucun lien de la compréhension vers la prosodie. Les scores aux dimensions de l'EMDF en CM1 sont prédits par les scores à l'EMDF en CE2, la compréhension écrite n'intervient pas.

5.2.3 Discussion

La figure 5.6 présente un schéma récapitulatif des relations longitudinales mise en évidence dans cette étude.

Du CE1 vers le CE2, les dimensions de décodage et de vitesse sont les meilleurs prédicteurs des dimensions de la fluence en CE2. Du CE2 au CM1, c'est le phrasé qui prédit le mieux les scores à ces dimensions en CM1. Il est intéressant de voir qu'en allant au delà du CE2, la prosodie prend un rôle plus important dans le développement des compétences en lecture.

On peut noter que les scores à chaque dimension sont liés aux scores de l'année précédente, excepté pour le décodage qui est uniquement prédit par d'autres compétences. Ces liens directs sont cependant assez limités car les scores de l'année précédente n'explique qu'une part, pafois trés faible, de la variance du score de l'année suivante (de 2 à 50% suivant les dimensions et les années). Comme dans les chapitres 3 et 4, on peut remarquer un changement entre le CE2 et le CM1. En effet, du CE1 vers le CE2, les différentes dimensions sont interdépendantes : l'expressivité est en partie prédite par le décodage et le phrasé par la vitesse. Les compétences d'automaticité semblent donc jouer un rôle dans l'acquisition des compétences prosodiques. Du CE2 vers le CM1, cette dépendance disparait, voir s'inverse. L'expressivité n'est plus prédite que par l'expressivité en CE2, et le phrasé par le phrasé en CE2. L'automaticité n'intervient plus dans le développement

des compétences prosodiques. On observe même que le décodage en CM1 est lié au phrasé en CE2, dénotant un impact des compétences prosodiques sur les compétences d'automaticité. Ces observations sont cohérentes avec l'impact de plus en plus important de la prosodie au fil des années d'apprentissage de la lecture.

Concernant les liens prosodie-compréhension, nos données ne nous permettent pas de mettre en évidence de lien bidirectionnel entre prosodie en lecture et compréhension écrite. On retrouve bien un lien de la prosodie vers la compréhension aussi bien du CE1 vers le CE2 que du CE2 vers le CM1. Toutes les dimensions prosodiques jouent un rôle dans cette relation : l'expressivité du CE1 vers le CE2, puis le phrasé du CE2 vers le CM1. Ainsi même si nous n'avons pas retrouvé de bidirectionnalité dans ce lien prosodie-compréhension, nous observons cependant qu'il s'agit bien d'un lien des dimensions prosodiques vers la compréhension, et que ce lien évolue dans le temps. Ainsi nous ne rejoignons pas les études publiées en néerlandais et en anglais. En effet, Veenendaal *et al.* (2016a) a montré un lien bidirectionnel du CM1 vers le CM2 puis du CM2 vers la 6e. De la même façon, Klauda et Guthrie (2008) se sont interessés à l'année de CM2 uniquement. Dans ces deux études, la direction compréhension vers prosodie est retrouvée chez les plus jeunes, puis évolue vers un lien prosodie vers compréhension. Les participants à notre étude ont été suivis du CE1 au CM1. Bien qu'ayant des participants plus jeunes que dans les études de Veenendaal *et al.* (2016a) et Klauda et Guthrie (2008), nous n'avons pas confirmé l'observation d'un lien de la compréhension vers la prosodie. Il est étonnant de retrouver uniquement le lien prosodie vers compréhension. Au vu de la littérature et de l'âge de nos participants, on s'attendrait à trouver plutôt le lien inverse : ceux qui comprennent sont ceux qui auront la lecture la plus expressive. A l'inverse ici, ceux qui sont le plus expressifs vont devenir ceux qui comprennent le mieux. Une explication possible est que les enfants ayant des bons scores d'expressivité et de phrasé en CE1 portent de l'attention à leur manière de lire et donc au texte, à sa structure et à sa signification, ce qui améliore leur compréhension. Les enfants portant moins d'attention à leur manière de lire à l'oral, et peut-être plus au décodage et à la vitesse à laquelle ils lisent, portent moins d'attention au sens du texte et ainsi développent moins leurs compétences de compréhension.

Un autre apport de cette étude est l'utilisation de trois tests de compréhension écrite, permettant d'évaluer trois formes de compréhension écrite différentes : globale, locale et en-ligne. Les tests généralement utilisés dans ce type d'étude sont plutôt des tests de compréhension globale, par exemple un texte suivi de questions, la plupart du temps des tests standardisés. Une seule étude s'est intéressée à un autre type de compréhension. Fernandes *et al.* (2017) ont pu lier la prosodie en lecture à la compréhension "sans effort", équivalente à ce que nous avons appelé compréhension en-ligne. Cependant nous ne retrouvons ici de lien qu'avec la compréhension locale. En effet, nous n'observons aucun prédicteur à la compréhension en-ligne, alors que celle-ci était liée au phrasé au sein d'un même niveau scolaire (cf chapitre 3). De la même manière, la compréhension globale semble ici uniquement prédite par la compréhension globale l'année précédente, alors qu'elle était également liée au phrasé au sein d'un même niveau. Ces résultats reposent donc la question du type d'évaluation utilisée pour la compréhension. Les tests utilisés ici pour la compréhension globale sont essentiellement prédits par les compétences en vocabulaire, laissant penser que les difficultés majeures de compréhension de ces textes sont liées au lexique. Il faut donc se poser la question des compétences réellement mesurées par ces tests de compréhension. Le manque de test standardisé en français est un frein. Concernant la compréhension locale, prédite par l'expressivité puis par le phrasé, il est important de souligner que nous n'avons ici pris en compte que le pourcentage de réponses correctes en lecture silencieuse, sans évaluer le temps de lecture, comme cela a été fait

en anglais. Nous nous concentrons ainsi sur la compréhension uniquement, sans interaction avec la fluence.

Finalement il pourrait être intéressant d'imaginer pour une future étude longitudinale sur un temps plus long, par exemple du CP au CM2, un test plus évolutif composé d'un texte à lire et d'un test de compréhension ciblé propre à chaque niveau scolaire. Ce suivi longitudinal plus long nous permettrait peut-être de faire apparaitre le lien bidirectionnel entre prosodie et compréhension et de confirmer l'importance croissante du phrasé et de l'expressivité dans les compétences en lecture quand les enfants grandissent et que celles-ci se développent.

Résultats majeurs — **Étude 7**

L'automaticité joue un rôle prédicteur important sur la fluence du CE1 vers le CE2. Du CE2 vers le CM1, c'est le phrasé qui prend ce rôle.

La compréhension en CE2 et en CM1 est prédite par les compétences prosodiques de l'année précédente : expressivité puis phrasé.

Nous avons mis en évidence un lien prosodie vers compréhension écrite variant avec le niveau et les compétences.

5.3 Conclusion

Dans ce dernier chapitre expérimental nous avons dans un premier temps présenté un nouvel outil d'évaluation automatique des scores à l'EMDF utilisant à la fois des paramètres verbaux et des marqueurs acoustiques, tout en prenant en compte la variabilité des productions prosodiques possibles. Nous avons montrer que cet outil est robuste, en obtenant de bons résultats sur différents textes. L'utilisation de cet outil avec nos données longitudinales a permis de montrer le développement de la prosodie obtenu sur les textes 1 et 2. On observe que le développement de la vitesse et du décodage est plus précoce que ceux du phrasé et de l'expressivité et que le rythme de développement de l'expressivité est plus lent que celui du phrasé et de l'automaticité. On rejoint donc les observations faites dans l'étdue 1. Le développement diffère dans les lectures des textes 1 et 2. Le texte 2 présente des scores plus bas, excepté pour l'expressivité, qui augmentent plus rapidement que le texte 1. Ces résultats viennent souligner l'importance du choix des textes dans l'évaluation de la prosodie. Des modèles liant les quatre dimensions de la fluence et la compréhension d'une année sur l'autre, ont permis de mettre en évidence l'importance croissante dans le temps de la prosodie dans le développement de la fluence et de la compréhension écrite. Il serait intéressant de renouveler cette étude dans un format plus long afin de confirmer nos hypothèses sur les tendances observées.

Chapitre 6

Discussion générale

6.1 Rappel des principaux résultats

6.1.1 Caractérisation du développement de la prosodie par évaluation subjective

L'EMDF, que nous avons traduit et adapté de la MDFS (Rasinski, 2004), est un échelle de mesure subjective de la fluence en lecture incluant quatre dimensions : deux dimensions prosodiques (expressivité et phrasé) et deux dimensions d'automaticité (décodage et vitesse). C'est un outil valide, fidèle et sensible du CE1 au CM1 avec les textes proposés. Du CM2 à la 5e, ces textes permettent seulement de détecter les élèves en difficulté, mais l'échelle peut-être utilisée avec d'autres textes plus complexes pour devenir un outil diagnostic. Avec un suivi strict du protocole d'utilisation, elle permet à des juges, même néophytes en matière de lecture d'enfant, d'évaluer de façon fiable les quatre dimensions de la fluence en lecture.

L'étude du corpus transversal avec cette échelle a montré un développement de chaque dimension trés rapide en CE1 et CE2, puis plus lent jusqu'en CM1 et plafonnant à partir du CM2. Le développement de l'expressivité se démarque de celui des autres dimensions : il est plus lent, avec des scores plus faibles. Ceux-ci plafonnent également, mais à des valeurs inférieures à celles obtenues par des lecteurs experts. On peut également noter que la variabilité entre les sujets est élevée en CE1 et CE2 et diminue fortement à partir du CM1, excepté pour l'expressivité. Cette étude va dans le sens d'un ordre dans l'acquisition des dimensions de la fluence, avec le décodage et la vitesse acquis plus précocément, le phrasé suivant et l'expressivité plus lente à acquérir. Nous ne pouvons cependant pas encore conclure sur le possible lien de causalité entre automaticité et prosodie à ce stade.

L'utilisation de l'EMDF dans l'étude 2 a permis de montrer que les scores de phrasé étaient corrélés à la compréhension écrite globale et en-ligne, sans interaction avec le niveau. Nous retrouvons donc dans cette étude le lien entre prosodie et compréhension déjà mis en évidence dans plusieurs langues, tout en ajoutant l'identification du phrasé comme principale dimension de la fluence corrélée à la compréhension.

6.1.2 Caractérisation du développement de la prosodie par les pauses et la respiration

La mesure de la prosodie se fait ici par des analyses acoustiques des pauses et de la respiration. Les pauses présentes dans le corpus transversal ont été annotées suivant leur caractère respiratoire ou non, grammatical ou non. Cette étude a repris les paramètres classiques présents dans la littérature : durée, nombre et grammaticalité des pauses, correspondant à un signe de ponctuation ou non. Ici, l'originalité est de quantifier les respirations, en utilisant à la fois leur marque acoustique sur les enregistrement audio et des ceintures respiratoires. Ce dispositif a permis d'identifier précisément les pauses respiratoires et de mesurer la durée entre inhalation et phonation, durée qui nous renseigne sur la coordination respiration-lecture et donc la planification des pauses.

L'étude de ces pauses a permis de retrouver les tendances observées en anglais et en espagnol (cf chapitre 1) : diminution du nombre et de la durée moyenne des pauses avec l'augmentation du niveau scolaire, diminution importante des pauses non grammaticales aprés le CE2. Cette étude a également permis de mettre en évidence la bimodalité de la durée des pauses dès le début de l'apprentissage. Ainsi les résultats connus ont pu être repris en séparant pauses courtes et longues, syntaxiques et respiratoires. La respiration est associé à 58 % des pauses longues quel que soit le niveau scolaire. Sa coordination avec la lecture se met en place dés le CE2 en s'appuyant sur la ponctuation. Les pauses sont alors encore nombreuses du fait d'une maturation de la coordination respiration-parole incomplète. Elle se rapproche des adultes dès le CM1. L'analyse de tous les paramètres étudiés montre un développement en quatre étapes : CE1 (peu de planification), CE2 (transition), CM1 à 6^{e} (plafonnement des performances s'approchant de celles des adultes), 5^{e} (planification similaire à celles des adultes). Un seul paramètre n'est pas mature en 5^{e}. La durée des pauses longues est en effet trés supérieure chez les adultes, comme chez les CE1. Les adultes, ici des lecteurs experts, utilisent probablement ces pauses de façon plus prononcée pour souligner la syntaxe, ce qui n'est pas encore le cas des 5es.

Dans l'étude 4, les résultats obtenus avec cette analyse acoustique ont été rapprochés des scores obtenus par les lectures à l'EMDF afin de déterminer quels paramètres étaient le plus suceptibles d'influencer la perception subjective de la fluence. Cette étude montre que le nombre et la durée des pauses ne sont que peu corrélés aux scores subjectifs. Il n'y a aucune corrélation avec le nombre de pauses non-grammaticales. Par contre, toutes les dimensions de l'EMDF sont corrélées au respect de la ponctuation. Les dimensions prosodiques sont essentiellement corrélées au respect de la syntaxe et de la ponctuation, y compris l'expressivité. On peut donc penser que les juges sont plus sensibles au respect des grandes frontières syntaxiques, qu'au nombre, placement et durée des pauses, dans leur perception de la fluence.

Finalement, l'étude 5 a mis en évidence la corrélation des scores de compréhension globale et en-ligne avec le nombre de pauses longues réalisées dans la lecture prosodique. Pour la compréhension en-ligne, le NMCLM intervient comme médiateur dans ce lien. Pour les enfants lisant lentement, un nombre de pauses longues élevé est plus révélateur d'une faible compréhension, que pour les enfants lisant plus rapidement. On peut noter que nous n'avons pas retrouvé de corrélation avec le nombre de pauses non grammaticales, comme c'était le cas dans d'autres langues.

6.1.3 Caractérisation du développement de la prosodie par un nouvel outil objectif et automatisé

L'outil présenté ici pour évaluer la prosodie permet d'obtenir un score à chaque dimension de la fluence en se basant sur des paramètres verbaux et acoustiques des lectures, analysés automa-

tiquement. La difficulté d'évaluation automatique des paramètres prosodiques réside dans le fait d'avoir une grande variabilité de productions expertes pour un même texte. Cet aspect a été pris en compte ici en utilisant les productions de multiples lecteurs experts pour construire des espaces de référence dans lesquels sont évaluées les lectures d'enfants. La combinaison de la prise en compte des paramètres verbaux et prosodiques des lectures permet d'obtenir une corrélation élevée entre scores donnés par l'outil automatisé et scores donnés par des juges humains. Cet outil est particulièrement adapté aux études longitudinales. Son objectivité permet de réaliser des évaluations multiples sans avoir à se soucier de la variabilité inter-évaluateur. Les scores aux évaluations répétées sont directement comparables. Sa précision permet de capter des variations fines dans les acquisitions prosodiques, uniquement mesurables avec des paramètres acoustiques.

Les lectures du corpus longitudinal du CE1 au CM1 ont donc été évaluées avec cet outil. Ces données ont permis de proposer des modèles de croissance pour chacune des dimensions de la fluence et pour les deux textes utilisés. Nous avons ainsi pu mettre en évidence une croissance régulière des scores aux dimensions phrasé et expressivité commune à tous les sujets étudiés. Le phrasé et la vitesse présentent une croissance qui diffère selon les sujets. Il est possible que cela soit dû au plafond qui semble être atteint à partir du CM1. Le taux de croissance est plus élevé pour le phrasé et la vitesse, et un peu moins pour le décodage et l'expressivité. L'expressivité semble se développer plus lentement et tardivement que les autres dimensions. On a également pu observer une influence du texte sur les scores et leur taux de croissance du CE1 au CM1. Le dialogue donne des scores plus faible mias augmentant de façon plus rapide, particulièrement pour l'expressivité et le décodage.

Pour finir, ces données ont permis également de proposer un modèle de développement des dimensions de la fluence du CE1 au CM1, en incluant la compréhension. Ce modèle a mis en évidence le lien de l'automaticité en CE1 vers la prosodie en CE2. Cela confirme donc la nécessité de la maitrise de l'automaticité dans le développement de la prosodie. Du CE2 vers le CM1, le phrasé a également un impact sur le décodage. La prosodie semble alors prendre le pas sur l'automaticité pour supporter le développement de la fluence. Les résultats avec la compréhension viennent appuyer cette observation. En effet, nous n'avons pas pu mettre en évidence de relation bidirectionnelle entre prosodie et compréhension. Cependant nous observons l'impact de la prosodie en CE1, puis en CE2 sur le développement de la compréhension. En CE1, l'expressivité est un bon prédicteur de la compréhension en CE2, puis le phrasé prend le relais du CE2 au CM1. L'influence de la prosodie sur la compréhension est donc bien présente dès le CE1 et l'expressivité et un phrasé approprié sont de bons indicateurs des compétences en compréhension.

6.2 Discussion des résultats

6.2.1 Evaluer et mesurer la prosodie

En mesurant la prosodie, nous introduisons une mesure qualitative de la lecture. Le NMCLM caractérise l'automaticité acquise par les enfants, mais ne semble pas être un indicateur suffisant des compétences en lecture des enfants. Il est nécessaire d'aller au-delà vers une évaluation plus large et qualitative de la lecture. Comme mentionné dans le chapitre 1, le choix du NMCLM en recherche, comme en éducation, relève du choix d'une mesure quantitative et simple à mettre en oeuvre. Evaluer qualitativement une lecture se révèle complexe, comme on a pu le voir tout au long

de ce travail. La difficulté majeure de la mesure de la prosodie est sa subjectivité. En effet, chaque lecteur a sa manière propre d'interpréter un texte et chaque auditeur son jugement subjectif. Il est donc difficile d'évaluer la prosodie et, à fortiori de mesurer une performance, sans objectif clair et quantifiable à atteindre.

L'évaluation de l'automaticité peut passer par le NMCLM. Mais une évaluation qualitative est également possible. L'EMDF proppose une évaluation subjective de l'automaticité à travers les dimensions de décodage et de vitesse. Dans l'étude 1, les scores aux dimensions automaticité de l'EMDF présentent une corrélation importante avec le NMCLM. Dans l'étude 6, les scores prédits sont liés aux paramètres verbaux : nombres d'erreurs et taux d'articulation. Mais dans une évaluation qualitative de la lecture, une part de subjectif apparait sur le décodage et la vitesse. Les juges n'evaluent pas de la même façon des erreurs de décodage rapidement corrigées, et des erreurs peu corrigées et nuisant à la compréhension du texte. De la même manière, la vitesse appropriée est aussi une notion subjective. en effet une lecture trop rapide nuit à la compréhension et à la qualité de la lecture. Cet aspect subjectif de la maitrise de l'automaticité est pris en compte dans l'EMDF par l'objectif du rythme conversationnel des lecteurs experts. Ces cotés subjectifs de l'automaticité ne sont pas captés par les paramètres verbaux directement. Cela apparait dans l'étude 4 avec l'influence du respect de la ponctuation et de la syntaxe sur les dimensions d'automaticité. Dans l'étude 6, les paramètres rythmiques apparaissent dans la prédiction des scores de décodage et de vitesse.

La dimension phrasé semble la dimension prosodique la plus simple à évaluer. Il est possible de comparer les productions des enfants aux adultes de références. En effet, le phrasé approprié est fortement contraint par la syntaxe et offre donc moins de possibilités de production que l'expressivité. Cela apparait sur la figure 5.2 de l'étude 6. On observe que l'espace de référence rythmique est plus restreint que le mélodique. Les enfants sont plus proches des productions adultes. Cependant l'étude 4 montre que le nombre, la durée et la grammaticalité des pauses ne sont pas suffisants pour caractériser un phrasé approprié. Le respect de la ponctuation joue également un rôle important dans l'appréciation subjective du phrasé. Dans le cas de mesures acoustiques de la prosodie, il faut donc choisir avec soin les paramètres suivis, afin qu'ils soient suffisamment révélateur de la qualité des lectures. Mais les paramètres acoustiques ne permettent pas l'interprétation des pauses. L'évaluation humaine, comme dans l'EMDF, est encore nécessaire pour interpréter les pauses et juger de leur adéquation avec le texte. On peut citer par exemple la durée des pauses longues. Ces pauses sont aussi longues chez les CE1 et les adultes experts mais pas pour les mêmes raisons. L'outil automatisé présente un compromis en utilisant l'espace de référence pour comparer les performances des enfants à celles de plusieurs adultes.

L'expressivité pose un problème plus complexe encore. Comme on l'a vu pour le phrasé, l'interprétation des pauses longues est essentielle dans l'évaluation de l'expressivité. Pour les variations de F0, il est difficile de comparer les contours mélodiques de la lecture d'un texte à une moyenne de courbe adulte. En effet, comme on peut le voir sur la figure 5.2 de l'étude 6, l'espace de référence des lecteurs experts pour l'expressivité est très large avec des productions trés variées. Dans cet espace, on observe trés peu de dispersion dans les lectures des enfants qui restent éloignées des lectures adultes. On peut également noter le caractère peu précis du but à atteindre dans l'échelle subjective pour l'expressivité. On demande au lecteur expert d'être expressif sans pour autant donner plus de détail. L'outil automatisé présenté ici semble être un bon compromis pour évaluer de manière objective, quantitative et reproductible l'expressivité d'une lecture. Il faut cependant s'affranchir de la dépendance au texte. En effet, il est plus difficile d'évaluer l'expressivité sur un

texte narratif que sur un dialogue. Nous reviendrons sur cette influence du texte dans le paragraphe 6.3

Pour conclure, le type d'évaluation le plus adapté va grandement dépendre de l'objectif à attendre. Pour une exploration des mécanismes cognitifs l'analyse des paramètres acoustiques est indispensable pour accéder aux détails de planification et de coordination des différents mécanismes nécessaires à la lecture. D'un point de vue éducatif, ou pour lier les compétences prosodiques en général à d'autres compétences, le passage par l'échelle subjective est suffisant et intéressant. L'utilisation d'un outil automatisé tel que celui présenté ici permet de s'affranchir de la variabilité liée aux juges et d'améliorer la précision et la reproductibilité des échelles subjectives et donc la comparaison de plusieurs textes, points de mesure, populations... Il est donc également plus facilement utilisable en recherche que les échelles subjectives humaines.

6.2.2 Caractérisation du développement de la prosodie

Les études 1, 3 et 6, mettant en évidence les étapes de développement de la prosodie, semblent s'accorder sur deux points essentiels à retenir ici : une étape importante dans le développement entre CE2 et CM1 et la particularité du développement de la dimension expressivité par rapport aux autres dimensions.

Les études en anglais et espagnol sur le développement de la prosodie se sont plutôt concentrées sur les toutes premières années CP et CE1, en anglais et CE2 pour espagnol. La synthèse du chapitre 1 faisait ressortir des compétences déjà bien installées en CE1 et trés proches des adultes en CE2. Dans les études présentées ici, nous avons pris le parti de scruter tous les niveaux scolaires du CE1 à la 5e. Dans les étude 1 et 3, on retrouve un développement trés rapide du CE1 au CE2 puis un ralentissement en CM1 et un plateau jusqu'au collège. Les performances sont alors proches des adultes experts mais demeurent en-deça pour la grande majorité des élèves, tout particulièrement dans la dimension expressivité.

Il semblerait donc que le développement de la fluence se joue principalement du CE1 au CM1. L'étude longitudinale, focalisant sur ces années, a confirmé une croissance régulière pour les dimensions prosodiques et qui semble ralentir pour les dimensions d'automaticité. De la même façon, elle a aussi permis de mettre en évidence l'inversion de causalité dans les différentes dimensions de la fluence. L'automaticité qui venait en support du développement des autres compétences jusqu'en CE2 cède la place à la prosodie, notamment au phrasé, à partir du CE2. On retrouve donc bien là aussi la charnière entre CE2 et CM1 à laquelle semble se produire une maturation des compétences prosodiques. Cette transition mériterait qu'on s'y intéresse plus spécifiquement, notamment au niveau de la maturation des différentes fonctions cognitives qui viennent en support de la prosodie : par exemple, l'acquisition de l'automaticité qui permet de libérer des ressources cognitives, la maturation de la respiration, de la planification, une meilleure compréhension de ce qu'on attend d'eux dans une lecture orale...

Le deuxième point remarquable et commun aux trois études de développement est le statut particulier de l'expressivité dans ce développement. Comme vu dans le paragraphe précédent, l'expressivité a tout d'abord un statut particulier de part son évaluation complexe, qui demande de faire des choix qualitatifs et subjectifs. Les jeunes lecteurs doivent également faire ces choix. En sont-ils conscients ? Comment choisissent-ils le ton à mettre dans leur lecture ? Quelle est l'importance d'avoir des modèles experts ? L'expressivité est peu enseignée, essentiellement implicitement

et par imitation. Pourtant celle-ci est importante pour la compréhension (cf étude 7 sur laquelle on reviendra dans le paragraphe suivant). Elle se développe plus lentement, et semble plafonner à des scores assez faibles, peu d'enfants atteignent des scores de lecteurs experts, même au collège. Dans l'outil automatique, de la même façon, la répartition des enfants dans l'espace des lecteurs experts est trés ramassée, loin d'atteindre la dispersion des adultes experts. Pourtant, pour les collègiens, le texte est facile, on le voit d'ailleurs sur l'étalonnage de l'EMDF avec un plafonnement des autres dimensions. Si quelques-uns atteignent le score maximal, ce n'est pas le cas de la majorité des lecteurs, particulièrement pour l'expressivité. Il serait intéressant de creuser plus avant cette stagnation. On retrouve cette différence avec les adultes experts également dans la durée des pauses longues qui reste inférieure chez les 5es, alors que tous les autres paramètres sont au niveau des adultes. On peut penser que cet écart est dû à un manque d'entrainement. Les collégiens ont peu l'ocasion de pratiquer la lecture à voix haute. D'autres hypothèses pourraient être également explorées : par exemple, un manque de motivation ou de confiance pour se mettre en scène.

6.2.3 Prosodie en lecture et compréhension écrite

Les trois études sur la compréhension ont permis d'une part de retrouver la corrélation entre prosodie et compréhension du CE1 à la 5^{e} mais également d'observer un lien directionnel de la prosodie vers la compréhension.

Nous avons retrouvé le lien fluence compréhension mentionné dans le chapitre 1. Au-delà de ce lien, nous avons observé l'implication du phrasé et du nombre de pauses longues, soulignant l'importance de l'attention à la syntaxe et de la planification des pauses. Dans l'étude longitudinale, nous avons observé une lien unidirectionnel de la prosodie vers la compréhension, dont les dimensions impliquées varient dans le temps. Ainsi notre étude vient ajouter un nouveau modèle, différent de ceux déjà proposés dans la littérature. Que ce soit au sein d'un même niveau ou en longitudinal, le phrasé ressort comme étant un indice important de la compréhension. L'expressivité apparait également chez les plus jeunes ce qui est étonnant mais rejoint l'étude de Miller et Schwanenflugel (2008) dans laquelle les variations de F0 en CP et CE1 prédisaientt la compréhension en CE2. L'expressivité apparaissait donc comme prédicteur de compréhension chez ces jeunes lecteurs. Dans la plupart des autres études, la compréhension chez les jeunes lecteurs est essentiellement liée au NMCLM (cf chapitre 1). Il est aussi intéressant de noter que nous avons controlé le NMCLM dans tous nos modèles de compréhension et que son effet n'est jamais significatif, alors que la prosodie explique une part de variance propre de la compréhension. Ces résultats complètent d'autres études remettant en cause le NMCLM comme bon indice de compréhension (par exemple Dowd et Bartlett 2019) et rejoignent les arguments appuyant la nécessité d'utiliser un autre outil plus qualitatif de type échelle subjective multidimensionnelle Rasinski *et al.* (2019)

Différents types de compréhension ont été utilisés ici : compréhension globale, en-ligne et locale. Les résultats sont congruents pour tous les types de compréhension : on observe un effet du phrasé et des pauses longues. On note cependant que les corrélations entre phrasé et compréhension globale et en-ligne au sein d'un même niveau, n'apparaissent pas dans les modèles longitudinaux. Dans ce dernier la compréhension en-ligne n'a aucun prédicteur. On peut donc se poser la question de ce qui est effectivement mesuré par ce test. Nous reviendrons sur ces questions dans le paragraphe suivant.

6.3 Limites

6.3.1 Dépendance aux textes

Le choix a été fait ici de travailler avec deux textes préconisés pour un niveau environ CE2, ne présentant pas de difficulté lexicale particulière et de garder le même texte pour tous les niveaux scolaires scrutés. Nous avons par ailleurs utilisé un texte narratif avec des phrases longues pour étudier le phrasé et un dialogue pour induire de l'expressivité. Il est apparu pendant les passations que le dialogue engendrait de fait plus d'expressivité que le texte narratif. Les enfants prennent plus facilement le rôle des personnages, ce qui n'est pas le cas dans le texte narratif pour lequel l'expressivité se dirige directement vers l'auditeur. L'étude 1 a en effet montré un décalage dans les scores en faveur du dialogue. L'évolution des scores est quand même similaire et la corrélation est forte entre les deux textes utilisés. Le choix des textes intervient donc quand on cherche à quantifier. Il apparait que cet effet de texte est également présent dans le développement avec des acquisitions plus rapide pour le dialogue, que pour le texte narratif, particulièrement pour les dimensions expressivité et vitesse. Il serait intéressant d'aller plus loin en intégrant également un texte informatif, comme l'ont fait Alvarez-Canizo *et al.* (2020); Benjamin et Schwanenflugel (2010). La lecture sera alors certainement moins expressive mais il serait intéressant de s'axer sur le phrasé et la compréhension dans des textes de plus en plus utilisés dans leur scolarité.

Le choix des textes est donc fondamental pour l'étude de la prosodie et particulièrement de son développement. Il est aussi possible d'utiliser des textes différents adaptés à chaque niveau scolaire. L'étude portera alors sur les capacités prosodiques sur des textes représentatifs de ceux communément rencontrés par les lecteurs. Cela nécessite cependant de disposer de textes de complexité étalonnée pour chaque niveau, ce qui n'est pas notre cas. La complexité d'un texte est difficilement quantifiable. Son évaluation demande la prise en compte de nombreux paramètres, y compris des paramètres subjectifs comme les connaissances préalables de chacun des lecteurs, la complexité des inférences ou la capacité du lecteur à interpréter les intentions de l'auteur...

On peut finalement se poser la question de la dépendance au texte de notre étape importante entre CE2 et CM1. Il serait aussi intéressant de proposer un texte plus simple aux plus jeunes, ne présentant aucun problème de décodage afin d'identifier si le phrasé et l'expressivité apparaissent plus précocément.

6.3.2 Préparation des lectures

Nous avons choisi pour toutes les études présentées ici de mesurer la prosodie sur la première lecture du texte, sans aucune préparation. Cela oblige le jeune lecteur à traiter simultanément le décodage, la compréhension et la performance orale, donc l'ajout de prosodie à sa lecture. C'est un exercice difficile car peu pratiqué. Les lectures expressives sont en général préparées et répétées. A minima, les enfants prennent connaissances du texte avant la performance orale. En faisant ce choix d'utiliser la première lecture, nous nous axons sur les mécanismes cognitifs mis en concurrence dans cette première lecture. Il est d'ailleurs probable que ce choix ait eu une influence sur le lien qu'on retrouve entre automaticité et prosodie. On voit que cette dernière se développe plus rapidement une fois l'automaticité acquise. Cet effet pourrait être dû à la concurrence des mécanismes cognitifs sous-jacents.

Une autre protocole possible serait de laisser les enfants découvrir le texte pour s'abstraire, au moins en partie de la contrainte de l'automaticité, et se centrer sur les compétences prosodiques. Ce type de protocole permettrait peut-être aux élèves les plus faibles de montrer de meilleures compétences prosodiques et changerait le profil de développement. Cependant ce type de protocole met également beaucoup d'autres compétences en jeu, comme la compréhension ou la mémoire. Il est possible qu'en permettant cette prélecture, nous retrouvions le lien directionnel compréhension vers prosodie. Il serait ainsi intéressant de mener les deux protocoles en parallèle pour explorer les différences prosodiques entre ces deux lectures et ainsi valider ou non un développement de la prosodie "préparée" diffèrent de la prosodie "spontanée".

6.3.3 Tests de compréhension

Le choix des tests de compréhension a déjà été partiellement commenté dans les discussions de certaines études. Il est cependant important et intéressant de revenir sur ce sujet car comme le choix des textes, il conditionne grandement les résultats.

Les textes avec questions, appelés ici compréhension globale, sont les tests de compréhension les plus utilisés dans la littérature pour rechercher les liens entre compréhension et prosodie. Cependant, il est difficile de savoir ce qui est exactement mesuré par ces textes. Pour des raisons pratiques, nous avons choisi de faire passer ces tests en groupe classe, sous la forme à laquelle les enfants sont habitués à l'école. Durant les passations, nous avons pu observer les élèves pendant ces tests. Les troubles de l'attention, le découragement face à la longueur d'un texte, la fatigue, le manque de motivation à aller chercher la réponse dans le texte, à l'écrire, l'envie de faire vite... Ces paramètres, particulièrement fréquents chez les plus jeunes, influent sur les résultats. On mesure, pour certains, plus l'attention et la motivation que les compétences en matière de compréhension. Dans les modèles que nous avons réalisés ici, nous avons d'ailleurs observé que le vocabulaire, variable controlée, explique la majorité de la variance des scores de compréhension globale. Ainsi cette mesure ne parait pas forcément la plus adaptée pour avoir une idée réelle et précise de la compréhension écrite des enfants dans sa globalité.

A partir du CE2 dans le suivi longitudinal, nous avons ainsi ajouté une mesure de la compréhension locale. Ce test, appelé Maze Test, est traditionnellement utilisé dans les pays anglo-saxons et représentatifs des compétences de compréhension (Parker *et al.*, 1992). Par rapport aux textes avec questions, son format est plus court et une seule lecture est nécessaire puisqu'il est complété au fur et à mesure. Il est donc plus attrayant, particulièrement pour les plus jeunes. Cependant cette mesure reste une mesure locale de la compréhension puisqu'il suffit de se référer à la phrase en cours ou à la précédente pour trouver la réponse correcte. Ce test ne permet pas de mesurer la compréhension globale et fine du texte. Dans le corpus transversal, nous ne l'avions pas proposé en CE1, par manque de temps dans le protocole. Nous n'avons donc pas de valeur en CE1 ce qui pourrait être une cause de l'absence de lien compréhension vers prosodie chez les plus jeunes.

Le troisième type de compréhension est la compréhension en-ligne. En comparant le NMCLM dans les textes signifiant et non-signifiant, nous avons voulu mettre en évidence l'apport de la compréhension à l'automaticité. La compréhension participe à l'anticipation du texte à venir et permet de lire plus rapidement un texte signifiant que non-signifiant. A priori, ce type de compréhension devrait être plutôt lié aux dimensions d'automaticité car la syntaxe est controlée et il n'y a aucune consigne d'expressivité. On constate cependant que le phrasé est corrélé à cette

compréhension en-ligne. Comme pour la compréhension globale, on peut se demander ce qui est vraiment mesuré. En effet, pendant les passations, nous avons constaté des stratégies différentes chez les enfants donnant des scores identiques, voir inverses à leurs compétences de compréhension. Sur le texte non-signifiant, certains lisent lentement en hésitant et répétant pour être sûrs de ne pas faire d'erreurs de décodage car ils ne peuvent s'appuyer sur le sens pour se corriger. A l'inverse, d'autres lisents rapidement sans se corriger car ils ne repèrent pas leurs erreurs. De la même façon, sur le texte signifiant, certains lisent plus lentement car ils cherchent à comprendre et à se corriger en cas d'erreurs, d'autres lisent rapidement car la compréhension les aide. Finalement, on observe aussi des enfants lisant les deux textes à la même vitesse, focalisés sur la vitesse sans aucun effet de la compréhension. Toutes ces stratégies vont donner des résultats identiques correspondant à des compétences différentes. Comme pour la compréhension globale, ces résultats ne sont pas forcément révélateur des compétences ciblées. Cela pourrait expliquer qu'aucun lien n'apparaisse, y compris avec les variables controlées, dans l'étude longitudinale, trop de stratégies différentes intervenant. Au vu de la corrélation entre phrasé et compréhension en-ligne au sein d'un même niveau scolaire, il serait intéressser d'étudier plus précisément le placement des pauses dans ces deux textes pour séparer l'impact de la compréhension et du traitement de la syntaxe dans la planification des pauses.

Pour finir, il est important de noter que nous avons ici mesuré la compréhension indépendamment de la prosodie en utilisant des textes différents, en lecture silencieuse et en autonomie, pour la compréhension globale et locale. Il serait intéressant de mesurer la compréhension du texte lu et enregistré pour mesurer la prosodie, comme l'ont fait Kocaarslan (2019). Nous explorerions alors un lien plus direct. Cela complexifie cependant la consigne car elle impliquerait de diriger la lecture à la fois sur la prosodie et la compréhension. Ce type de mesure de la compréhension pourrait être utilisé dans le cadre d'un protocole avec découverte du texte. On pourrait imaginer une première lecture ayant pour consigne la compréhension, suivie de questions à l'oral, avec l'expérimentateur, puis une lecture du même texte ayant pour consigne la performance orale et l'expressivité. La mesure de la compréhension écrite en passation individuelle permettrait d'avoir une meilleure mesure de la compréhension globale en s'affranchissant des perturbations dues à l'attention, la motivation ou la difficulté à formuler les réponses.

6.3.4 Diversité des sujets et des stratégies

Comme mentionnés dans les paragraphes précédents nous avons ici mesuré des scores, prosodie, automaticité, pauses..., sans aucune interprétation qualitative de ceux-ci. Nous avons décrit le développement général de la prosodie. Il est cependant important de préciser que les étapes d'acquisition que nous avons observées et définies ici recouvrent de nombreuses stratégies différentes, que ce soit de lecture ou de compréhension. Nous avons abordé dans le paragraphe précédent les stratégies différentes couvertes par des résultats de compréhension en-ligne. Nous avons également évoqué l'absence d'interprétation des pauses dans notre étude de celle-ci. Le meilleur exemple est la durée des pauses longues identiques pour les CE1 et les adultes, sans pour autant pouvoir l'interpréter de la même façon. Pour l'outil automatisé, l'expressivité est comparé à celle des adultes, mais sans regard sur l'adéquation de celle-ci avec le texte. Ainsi il serait intéressant d'analyser plus finement les paramètres prosodiques par une analyse qualitative humaine. Annoter précisément les fonctions des pauses (hésitation, décodage, phrasé, respiration) permettrait de mieux interpréter les stratégies cognitives associées et leur lien avec la planification et les autres compétences en lecture. De la même façon, il serait intéressant d'aller plus loin dans l'analyse de l'expressivité en la liant à

la compréhension fine du texte : quelles émotions pour quel personnage... Une telle analyse, mise en lien avec un protocole de test de la compréhension du texte lu avec expressivité, permettrait de mieux comprendre la relation entre prosodie et compréhension. Ce protocole pourrait permettre de faire apparaitre le possible lien directionnel que nous n'avons pas observé dans nos données.

Une autre question se pose : la prise en compte du bilinguisme. Tout profil compris (langue maternelle et seconde) 26% des enfants enregistrés étaient bilingues. Quelle est l'influence de ce bilinguisme sur le développement de la prosodie en lecture ? Nous n'avons pas cherché dans nos études à identifier un éventuel effet du bilinguisme (français langue maternelle) sur les paramètres étudiés. Dowd et Bartlett (2019) ont montré que les enfants bilingues lisent moins vite pour une compréhension égale aux enfants monolingues. Il est possible qu'il y ait également un effet du bilinguisme sur la prosodie en lecture. Notamment si les enfants sont soumis à des modèles de lecture orale dans une autre langue. Il serait intéressant d'explorer cet effet du bilinguisme dans nos modèles pour observer d'éventuelles différences entre enfants bilingues et monolingues. De plus, les enfants ayant appris le français à l'école, ne le parlant pas à la maison, sont moins soumis aux modèles de lecture en français. Comment alors évaluer le développement de leur prosodie en lecture ? Cette question est d'autant plus importante qu'elle concerne un grand nombre d'enfants dans les écoles françaises.

La dernière question est l'influence que la réalisation de l'étude elle-même a pu avoir sur le développement de la prosodie en lecture chez les enfants suivis dans l'étude longitudinale. En effet, la majorité des enseignants dont les classes ont participé à l'étude, se sont interessés au sujet et y ont été sensibilisés. Les enfants ont compris ce qu'on attendait d'eux et ont peut-être également été plus attentifs à leur façon de lire à l'oral. Il est donc possible que leur trajectoire de développement soit un peu plus rapide que celle qu'on pourrait observer sans intervention extérieure. Cependant l'étape entre CE2 et CM1 apparait dans le corpus transversal et le corpus longitudinal. Cet effet est donc relativement robuste puisque présent dans nos deux populations.

6.4 Perspectives

6.4.1 En recherche

Evaluation de la prosodie

Concernant l'évaluation de la prosodie, plusieurs grands axes sont encore à explorer. Tout d'abord, il serait intéressant de prolonger la caractérisation des marqueurs acoustiques. Comme évoqué précédemment, cela passe pas une meilleure catégorisation des fonctions des pauses et par l'analyse acoustique des marqueurs mélodiques et d'intensité. En effet, nous n'avons pas abordés ici les variations mélodiques liées à l'intonation et à l'expressivité d'un point de vue acoustique. Cette caractérisation est en effet complexe car nous ne pouvons pas les comparer à une seule production adulte comme dans le cas des pauses. Pourtant, comme vu pour les pauses, ces marqueurs acoustiques sont précieux pour renseigner les mécanismes cognitifs sous-jacents aux compétences de lecture. Il est donc nécessaire de prolonger l'étude des marqueurs acoustiques par les marqueurs mélodiques : montée et descente de voix en fin de phrase, accentuation de mot-cible dans le dialogue... Ces mécanismes font appel à l'anticipation sur la syntaxe et le sens du texte et nous apporterait donc plus d'informations sur le développement de la planification de la prosodie.

Nous avons ici uniquement exploré le lien entre pauses et perception de la fluence par les juges utilisant l'EMDF. Il serait maintenant intéressant de continuer à caractériser les paramètres influençant la perception dans l'évaluation subjective. On pourrait en effet ajouter les variations de F0, que nous n'avons pas inclus ici, dans les modèles de l'étude 4. Les changements de registre sont également un point à explorer. Il est en effet possible que les juges humains soient influencés par l'âge supposé des enfants qu'ils entendent. On aura tendance à juger avec plus de bienveillance des trés jeunes lecteurs que des lecteurs plus expérimentés. Le jugement subjectif de l'âge supposé, voix d'enfant ou voix de pré-adolescent, pourrait ainsi avoir une influence sur le score final.

Une autre grosse étape dans le développement de l'outil automatique est l'amélioration de l'alignement automatique du texte sur la voix des enfants en vue des analyses. En effet la prise en compte des nombreuses erreurs, répétitions, omissions, etc est complexe à traiter automatiquement. A l'heure actuelle l'alignement automatique des lectures d'enfants nécessite une vérification manuelle rigoureuse. Comme l'outil automatisé et l'échelle subjective, ses performances sont dépendantes du texte. Il est donc également nécessaire de continuer à développer ces outils.

Cognition et développement

Le travail présenté ici a permis de mettre en évidence l'intérêt de la prise en compte de la prosodie en lecture et a donné les grandes étapes de son développement d'un point de vue qualitatif et cognitif. Cependant de nombreuses questions restent en suspend en matière de mécanismes cognitifs mis en place dans son développement et à fortiori dans son lien avec la compréhension. Nous nous sommes concentrés dans ce travail sur la description du développement de la prosodie. Nous avons lié ce développement à la compréhension et à l'automaticité en controlant quelques paramètres influençant également le développement de la compréhension, comme le vocabulaire, le raisonnement non-verbal et la compréhension orale. Au delà de l'automaticité, d'autres paramètres sont suceptibles d'influencer le développement des compétences prosodiques.

Pour poursuivre les travaux présentés ici, il serait intéressant d'aller plus loin dans l'exploration des mécanismes de planification. Nous nous sommes concentrés sur la planification des pauses, en utilisant un texte narratif. L'analyse de l'intonation dans le texte 2 des corpus transversal et longitudinal permettrait d'avoir des informations sur l'anticipation de l'intonation en fin de phrase. En espagnol, malgré une ponctuation des phrases interrogatives en début de phrase, l'intonation spécifique à ces phrases ne se développe pas plus tôt qu'en anglais (Álvarez-Cañizo *et al.*, 2015). Il semblerait donc que l'aide de la ponctuation n'est pas suffisante pour de jeunes lecteurs qui ne maitrisent peut-être pas encore ce type d'intonation. L'analyse de notre corpus en français permettrait d'apporter de nouvelles données sur ce sujet dans une troisième langue et de confirmer ou non cette hypothèse. Afin de mieux comprendre les mécanismes permettant l'anticipation des pauses syntaxiques et des fins de phrases, il serait intéressant de mener une étude en oculométrie du même type que celle d'Alamargot *et al.* (2010). Dans cette étude les auteurs utilisent la captation simultanée des mouvements de stylo et du regard pendant une production d'écrit. Ils ont ainsi pu montrer des stratégies différentes suivant l'âge des participants à travers l'observation des pauses dans le geste d'écriture. On pourrait transposer cette technique à notre cas. On peut en effet supposer que la lecture silencieuse est menée suffisamment en amont de la production orale pour identifier la structure syntaxique et la ponctuation et ainsi ajuster la planification des pauses. Une telle étude chez des enfants de CE1, chez qui les pauses sont peu planifiées, de CM1 ou CM2, chez qui la planification est déjà en place et chez des adultes experts, permettrait de faire le lien

entre anticipation visuelle et planification. L'anticipation visuelle peut-elle expliquer les différences de productions prosodiques ? Si c'est le cas, une étude longitudinale du CE1 au CM1 explorant l'anticipation en oculométrie permettrait de comprendre sa mise en place. Dans un deuxième temps, il serait intéressant d'étendre une telle étude à l'expressivité et à l'anticipation des sentiments d'un personnage dans un dialogue comme le texte 2 utilisé ici. La caractérisation de la planification de la prosodie pourrait également permettre de lier cette compétences aux autres fonctions éxécutives également nécessaires à la compréhension.

Le développement des dimensions prosodiques de la fluence sont liées au développement de l'automaticité. Il est trés probable que l'automaticité ne soit pas le seul prérequis nécessaire. Ainsi il serait intéressant d'aller plus loin dans la recherche des prérequis et précurseurs de ce développement. Plusieurs points sont alors à explorer. Le premier axe à explorer est le lien avec les contributeurs traditionnels des compétences en lecture, par exemple vocabulaire, mémoire de travail, conscience phonologique et syntaxique... Au delà de ces prédicteurs potentiels, on sait que toutes les figures prosodiques, notamment pragmatiques, comme l'ironie, sont encore en cours de développement quand les enfants apprennent à lire en primaire (Wells *et al.*, 2004) Celles-ci sont également liées aux capacités de perception de la prosodie, que ce soit le rythme de la parole ou ses variations mélodiques (Holliman *et al.*, 2017). Une étude transversale du CP au CM2 pourrait permettre de mettre en lien perception et production de la prosodie en parole et en lecture. De la même façon, une étude longitudinale pourrait être menée, peut-être avec des enfants plus jeunes, en mesurant à la fois les capacités de perception et de production de la prosodie en parole et le lien causal potentiel avec le développement de la prosodie en lecture.

Finalement, comme mentionné dans les limites, nous avons suivi ici un public tout-venant, aboutissant à une caractérisation générale du développement de la prosodie en lecture. Il serait maintenant intéressant d'aller vers des publics plus particuliers qui pourraient également nous renseigner sur les mécanismes de développement. Certaines études existent en français et en espagnol (Lalain *et al.*, 2014; Suárez-Coalla *et al.*, 2016) montrant les caractéristiques particulières des dimensions prosodiques en lecture orale des personnes dyslexiques et ayant des troubles du langage : plus de pauses, plus longues... L'exploration de ces particularités rythmiques, en ajoutant peut-être des tests de perception de la prosodie, pourrait être mis en lien avec la compréhension et la planification des pauses. De plus, il n'existe aucune étude sur le public bilingue qu'il soit précoce ou tardif. Une étude comparative du développement de la prosodie dans deux langues permettrait d'explorer des différences potentielles de développement dues à la langue, par exemple entre des langues de transparence différente. Ainsi l'étude du développement de la prosodie en lecture chez ces lecteurs particuliers permettrait d'approfondir les connaissances sur ses mécanismes cognitifs et de développement.

Pour conclure, ces études, complémentaires à celles menées dans ce travail de thèse, permettraient de mieux caractériser le développement de la prosodie en lecture et ses prérequis et donc de mieux le prendre en compte dans l'enseignement de la lecture.

Prosodie et compréhension

Nous avons montré ici la corrélation entre compréhension et prosodie en lecture et le lien causal entre compétences en expressivité et phrasé et la compréhension écrite. Ce lien de causalité a été observé dans l'étude longitudinale. Nous avions alors fait l'hypothèse que seul le lien prosodie vers

compréhension apparaissait dans cette étude car les enfants attentifs à mettre de l'expressivité dans leur lecture étaient plus attentifs au contenu du texte et donc développaient plus leur compréhension des textes lus. Un entrainement explicite à la lecture prosodique pourrait permettre de tester cette hypothèse. En effet, la lecture orale expressive est peu entrainée et explicitement enseignée. La mise en place d'un programme d'entrainement à la lecture expressive permettrait de mesurer son impact sur l'acquisition des compétences en lecture. La mesure de la compréhension écrite avant et aprés cet entrainement, peut-être également à distance de celui-ci, pourrait permettre de vérifier que l'attention portée à l'expressivité en lecture orale permet d'améliorer la compréhension écrite.

On peut alors se poser la question du type d'entrainement à mettre en place. En effet, celui-ci pourrait se limiter à la mise en contact avec des modèles de lecteurs experts à reproduire, et une visualisation explicite des pauses et de l'intonation à prévoir pour les plus jeunes. Ce type d'entrainement entrainerait les bases prosodiques nécessaires à la compréhension d'un texte : pauses syntaxiques et intonation appropriée. Nous avons utilisé ce type d'entrainement dans le cadre du projet Fluence, avec un karaoké de lecture utilisant le paradigme de lecture synchrone (Godde *et al.*, 2017a). Ce type d'entrainement correspond plutôt à une compréhension littérale du texte. Au delà de ces compétences nécessaires à la compréhension, l'entrainement de la prosodie peut également aller plus loin. L'expressivité passe en effet par l'expression d'une ambiance, d'émotions et d'intention des personnages voulues par l'auteur. Les lecteurs experts sont capables d'exprimer tout cela dans leur lecture en jouant à la fois sur le rythme, la mélodie et l'intensité de leur voix pendant la lecture. Ces compétences sont intimement liées à la compréhension fine du texte, à l'inférence des non-dits et à l'attribution d'intention à l'auteur et aux personnages. On retrouve aussi les entrainements de type lecture théatralisée (Young et Rasinski, 2018). Travailler sur cet aspect de la mise en voix d'un texte nécessite donc de travailler explicitement sur la compréhension de celui-ci mais en ayant un but autre que la pure compréhension. En effet, il s'agit alors de faire passer par la voix toutes les intentions de l'auteur afin de faciliter la compréhension fine des auditeurs. Ce type d'entrainement aurait donc nécessairement un impact sur la compréhension. Les différents types d'entrainement utilisables à cet effet seront développés dans le paragraphe suivant.

Dans ce travail nous nous sommes intéressés uniquement à des textes narratifs auxquels sont principalement exposés les enfants de primaire en lecture orale. Cependant en avançant dans les niveaux scolaires, ceux-ci sont de plus en plus exposés à des textes explicatifs dans lesquels la lecture orale est moins présente, mais la compréhension nécessaire à la recherche d'informations. Il existe peu d'étude sur la prosodie dans les textes documentaires. On peut imaginer que l'expressivité jouera un rôle moindre dans ce type de texte. Par contre, le phrasé et la mise en valeur des frontières syntaxiques et des relations de dépendances entre constituants peut permettre une meilleure compréhension des phrases complexes. En effet, plusieurs études ont montré de l'emphase dans ces paramètres quand les lecteurs sont confrontés à des textes plus complexes, y compris chez les adultes (Dodane et Brunellière, 2006; Kentner et Vasishth, 2016). Une étude de type acoustique sur le placement des pauses par rapport à la syntaxe du texte et la compréhension directe de celui-ci pourrait également nous renseigner sur le lien entre prosodie et compréhension. Le même type de protocole avec plusieurs textes de difficulté croissante pourrait permettre de mettre en évidence une mise en relief de la prosodie : pauses plus longues, variations d'intonation plus amples, allongements de syllabes plus marqués... Il serait intéressant de mener ce type d'études avec des sujets plus agés, collégiens et lycéens, confrontés à des textes explicatifs de plus en plus complexes.

Finalement, une grande part des textes lus par les adolescents et les adultes le sont en lecture silencieuse. Des études ont montré par l'observation du temps de fixation que la prosodie est éga-

lement présente en lecture silencieuse, particulièrement pour désambiguiser des phrases complexes (Dodane et Brunellière, 2006). Là encore, on retrouve le lien avec l'aide à la compréhension. Cet aspect n'a, à notre connaissance pas été étudié chez les jeunes lecteurs. On observe fréquemment les plus jeunes lire à voix basse quand ils ont besoin de comprendre, comme si d'entendre le texte aidait à sa compréhension. A quel moment se mettent en place ces mécanismes de prosodie en lecture silencieuse ? Ceux-ci varient-ils avec la complexité des textes ? L'exploration de cet aspect pourrait permettre de mieux comprendre le lien entre prosodie et compréhension et donc également de mieux préparer les jeunes lecteurs à lire des textes complexes en lecture silencieuse.

Les quelques études causales sur la prosodie et la compréhension ont permis de mettre en avant la complexité de ce lien et l'importance de l'explorer. La poursuite et l'approfondissement de son étude par les pistes proposées ici permettraient de mieux le comprendre et donc de mieux enseigner la compréhension et prendre en charge les lecteurs en difficulté.

6.4.2 Enseignement de la lecture et application en classe

La caractérisation du développement de la prosodie en lecture, de son lien avec la compréhension écrite, améne à proner une approche plus qualitative de l'enseignement et de l'évaluation de la lecture chez les jeunes lecteurs. Rasinski *et al.* (2019) utilise dans sa conférence le terme d'"art de la fluence". Il présente une étude sur le travail théâtral et l'impact positif que celui-ci peut avoir sur la lecture, orale et silencieuse, particulièrement chez les lecteurs en difficultés. Travailler la fluence comme un art engage les apprenants au delà de l'apprentissage. Cela engage performance et émotions dans la lecture. La lecture, les répétitions, l'expressivité deviennent alors un jeu et non plus uniquement un apprentissage ardu. Ce travail de la lecture qualitative comme un art permet donc l'acquisition de compétence diverses : prosodie, compréhension, fluence, etc, en gardant entière la motivation de l'apprentissage.

Cela pose également la question de l'enseignement explicite de la prosodie, et ce dés le début de l'enseignement de la lecture. Au vu de nos résultats, cet enseignement explicite pourrait être particulièrement efficace du CE1 au CM1. Mais sensibiliser les plus jeunes dés le CP, voir la maternelle, permettrait une acquisition comme une compétence indispensable à la lecture et non comme une compétence accessoire à l'automaticité et à la compréhension. Comme nous l'avons déjà évoqué, hors respect de la ponctuation, le respect de la syntaxe et l'expressivité est peu expliqués aux apprentis lecteurs. Dans les protocoles exposés par Rasinski *et al.* (2019)), comme la poésie et le théatre, on explicite où et quand respirer, où élever la voix, comment faire passer telle ou telle émotion et pourquoi.

Il existe différent type d'entrainement explicite de la prosodie. Certains entrainement à la fluence permettent d'agir sur la prosodie : lecture répétée (Chard *et al.*, 2002; Kuhn et Stahl, 2003) à la base de nombreuses méthodes d'entrainement, variété et nombre de textes accrus (Kuhn, 2005; Allington, 2014), lecture en duo avec un modèle expert (Chard *et al.*, 2002). D'autres méthodes sont centrées sur la prosodie comme la lecture théatralisée (Young *et al.*, 2016; Young et Rasinski, 2018), l'entrainement musical, particulièrement rythmique (Young *et al.*, 2016; Long, 2014), la lecture répétée centrée sur la prosodie (et non sur l'automaticité) (Calet *et al.*, 2017; Young *et al.*, 1996), indiçage du texte par des marquages des pauses et de l'intonation (Levasseur *et al.*, 2008; Patel *et al.*, 2014) et la lecture synchrone dans laquelle le lecteur lit en même temps qu'en lecteur expert en suivant son modèle (Patel *et al.*, 2014; Godde *et al.*, 2017a). La combinaison de ces différents

types d'entrainement permet de multiplier les approches explicites, de modélisation et de répétitions pour augmenter les effets de l'entrainement sur la prosodie et la compréhension. Par exemple, le karaoké de lecture utilisé dans le projet Fluence utilise à la fois les paradigmes d'indiçage, pour indiquer le placement des pauses et de la respiration, et de lecture synchrone pour proposer un modèle d'expressivité, en lecture répétée (Godde *et al.*, 2017a). L'effet de ces entrainements sur la compréhension n'a pas toujours été testé. On trouve cet effet avec la lecture répétée prosodique, la lecture théatralisée et l'entrainement musical. Le test de l'effet d'un entrainement combinant plusieurs méthodes, comme le karaoké, donnerait des informations supplémentaires sur le lien prosodie-compréhension.

On pourrait également faire le lien entre l'enseignement explicite de la prosodie et l'enseignement de la musique. En effet nous avons pu constater lors des passations que les élèves musiciens, en particulier chanteurs et instrumentistes à vent, sont beaucoup plus conscients et attentifs à leur respiration et pour les plus grands aux nuances apportées à leur voix, que des non-musiciens. En musique, l'enseignement de la respiration et des nuances est explicite et les indications sont notées sur la partition. Il serait donc trés intéressant d'étudier spécifiquement le développement de la prosodie chez ces apprentis lecteurs musiciens. Ont-ils une meilleure expressivité et un meilleure respect du phrasé que d'autres enfants non musiciens de même niveau d'automaticité ? On pourrait également imaginer un travail et une explicitation du même type que les annotations d'une partition en annotant les pauses à faire dans un texte et les variations mélodiques d'intonation, voir d'expressivité. De la même manière, pour les enfants suffisamment lecteurs travailler la lecture de texte avec un modèle prosodique et des annotations de pauses, en répétition ou en simultané, pourrait améliorer leur prise en compte dans leur lecture orale, puis par généralisation dans la lecture silencieuse.

En effet, entendre régulièrement des lecteurs experts pour avoir une idée du résultat final attendu, de ce qu'est une lecture expressive, est déjà une forme d'enseignement explicite. Les enfants de maternelle, pré-lecteurs, sont déjà capable dans leur jeu d'imitation de jouer à lire des livres en imitant la prosodie adulte spécifique à la lecture. D'un point de vue social, il serait donc intéressant d'étudier l'effet de l'exposition des enfants à la maison et à l'école, en prenant en compte la lecture orale des parents et des enseignants. L'exposition en classe à des modèles variés de lecture orale est mise en place dés la maternelle : lecture offerte par la maitresse, écoute de livres audio lus par des acteurs, participation à des spectacles de contes,... Pour les plus grands cette exposition peut être mise en place sous forme de lecture en duo avec les parents ou des lecteurs plus confirmés. Il existe de multiple formes d'exposition à des modèles de lecteurs experts. L'effet de l'exposition à la maison serait également intéressant à explorer. Cela pourrait prendre la forme de programme de lecture orale quotidienne avec les parents. La quantification de l'impact sur la prosodie en lecture, et sur la motivation à lire de façon expressive de ces types d'exposition serait intéressant à analyser et apporterait des pistes quant à l'effet de l'explicitation de l'expressivité.

Finalement, nous avons présenté dans les chapitres 3 et 5 des outils d'évaluation utilisables en classe pour le suivi des acquisitions des élèves. Pour commencer l'utilisation de l'échelle subjective serait intéressante à tester directement en utilisation par les élèves. Dans une optique d'évaluation l'EMDF permettra aux enseignants de disposer d'un outil permettant une approche plus qualitative de la lecture. L'outil automatique, une fois diffusé, pourrait permettre un suivi assez fin des acquisitions des élèves tout au long de l'année scolaire et même au niveau du cycle. En effet, la mesure étant effectuée toujours par le même outil calibré permet de suivre les acquisitions de manière qualitative même en changeant d'enseignant. Cela pourra permettre le suivi sur le long terme notamment d'élèves en difficulté.

L'EMDF a pour l'instant été soumise aux enseignants comme outil d'évaluation et l'outil automatique est encore en phase de calibrage et de tests. Mais ces deux outils explicitent ce qu'est une lecture orale experte et présentent donc également un intérêt formatif pour l'élève. L'utilisation directe de l'EMDF par les élèves leur permettrait de mieux comprendre quels types de compétences sont attendus d'eux et ainsi de mieux diriger leurs apprentissages. Prenons par exemple la vitesse de lecture. Afin d'atteindre un niveau expert, il est nécessaire de lire à un rythme conversationnel et non le plus vite possible. Cette nuance est difficile à appréhender par certains jeunes lecteurs. La compétence clairement défini de "rythme conversationnel" pourrait les aider à faire plus attention à leur vitesse et ralentir pour faciliter la compréhension de l'auditeur. Notre outil automatisé permet pour l'instant de donner un score à chaque dimension par comparaison avec des références expertes. Il serait intéressant de développer un feedback associé à ce score en précisant quels paramètres font augmenter les scores et lesquels le font diminuer. On peut ainsi imaginer une évaluation formative automatique et autonome pour les enfants. La lecture d'un texte serait enregistrée, analysée par l'outil automatique qui leur renverrait les points forts et les points à améliorer pour augmenter leur score. Ce feedback immédiat permettrait aux élèves de répéter plusieurs fois une lecture en jouant sur la prosodie. Le côté autonome permet également de s'affranchir du stress de l'évaluation devant un enseignant. Le feedback automatique et répété, impossible pour l'enseignant à mener en classe avec chaque élève, leur permettrait de se rendre compte de leur progression, d'identifier les points à travailler et de rester motivé.

Les travaux menés dans cette thèse incitent, je l'espère, à une meilleure prise en compte de la prosodie dans l'enseignement de la lecture et ouvrent la porte à des applications directes pour travailler cette compétence avec les élèves.

Conclusion

Cette thèse avait pour objectif d'aborder le développement de la prosodie en lecture d'un point de vue de la mesure et de la cognition. Nous souhaitions mettre en évidence les étapes de développement de la prosodie en lecture chez les apprentis lecteurs du CE1 à la 5e. Nous voulions également observer la mise en place de la planification des pauses et de sa coordination avec la respiration. Finalement nous voulions explorer les liens entre la prosodie, l'automaticité et la compréhension et leur évolution au cours du développement. Pour chacun de ces objectifs, nous avons souhaité proposer un outil de mesure de la prosodie en lecture, prenant en compte sa complexité, et permettant d'obtenir des données fiables et représentatives de nos participants.

Nos travaux ont permis de mettre en évidence les grandes étapes de développement de la prosodie en lecture, notamment en matière de planification des pauses et de la respiration. Nous avons observé en particulier le développement rapide du CE1 au CM1, avec une maturation des compétences entre CE2 et CM1. Si les productions au collège sont trés proches des lecteurs adultes, l'expressivité n'est cependant pas encore complètement acquise. L'expressivité est en effet apparue comme une compétence à part, se développant plus lentement et tardivement que l'automaticité et le phrasé. Nous avons également pu montré que le développement des compétences prosodiques est lié à celui de l'automaticité et de la compréhension. L'expressivité et le phrasé semble être de bons indicateurs de la compréhension écrite. Finalement, nous avons aussi proposé trois méthodes de mesures permettant de mesurer de façon simple, fidèle et sensible la prosodie dans les lectures d'enfants. L'échelle subjective permet une évaluation rapide et globale des compétences en fluence, y compris la prosodie et est utilisable facilement en classe. L'analyse objective de marqueurs acoustiques permet une exploration plus fine des mécanisme cognitifs par leur analyse. Finalement, l'outil automatisé proposé, prenant en compte la variabilité des productions possibles, basé sur l'échelle subjective et des analyses acoustiques, permet un suivi fin des compétences dans le temps.

Pour conlure, ces travaux montrent, je l'espère, la nécessité de s'intéresser au développement et à l'apprentissage de la lecture à voix haute dans l'enseignement de la lecture, et proposent des outils permettant de le faire. En effet, en travail du phrasé et de l'expressivité présente des bénéfices pour l'apprentissage de la lecture et de la compréhension. Ces considérations éducatives et l'intérêt de prendre en compte la prosodie dans l'apprentissage de la lecture ne doivent cependant pas faire oublier que celle ci est aussi un art résidant dans le plaisir du lecteur à interpréter un texte et un plaisir pour l'auditeur qui en écoute la musique.

Bibliographie

Alamargot, D., Plane, S., Lambert, E. et Chesnet, D. (2010). Using eye and pen movements to trace the development of writing expertise : case studies of a 7th, 9th and 12th grader, graduate student, and professional writer. *Reading and Writing*, 23(7):853–888.

Allington, R. L. (2014). How reading volume affects both reading fluency and reading achievement. *International Electronic Journal of Elementary Education*, 7(1):13–26.

Alvarez-Canizo, M., Cueva, E., Cuetos, F. et Suarez-Coalla, P. (2020). Reading fluency and reading comprehension in spanish secondary students. *Psicothema*, 32(1):75–83.

Álvarez-Cañizo, M., Martínez-García, C., Cuetos, F. et Suárez-Coalla, P. (2019). Development of reading prosody in school-age spanish children : a longitudinal study. *Journal of Research in Reading*.

Álvarez-Cañizo, M., Suárez-Coalla, P. et Cuetos, F. (2015). The role of reading fluency in children's text comprehension. *Frontiers in psychology*, 6:1810.

Álvarez-Cañizo, M., Suárez-Coalla, P. et Cuetos, F. (2018). Reading prosody development in spanish children. *Reading and Writing*, 31(1):35–52.

Applegate, M. D., Applegate, A. J. et Modla, V. B. (2009). "she's my best reader ; she just can't comprehend" : Studying the relationship between fluency and comprehension. *The reading teacher*, 62(6):512–521.

Arcand, M.-S., Dion, E., Lemire-Théberge, L., Guay, M.-H., Barrette, A., Gagnon, V., Caron, P.-O. et Fuchs, D. (2014). Segmenting Texts Into Meaningful Word Groups : Beginning Readers' Prosody and Comprehension. *Scientific Studies of Reading*, 18:208–223.

Bailly, G. et Gouvernayre, C. (2012). Pauses and respiratory markers of the structure of book reading. *In Interspeech*, pages 2218–2221.

Bailly, G., Raidt, S. et Elisei, F. (2010). Gaze, conversational agents and face-to-face communication. *Speech Communication*, 52(6):598–612.

Bailly, G., Rochet-Capellan, A. et Vilain, C. (2013). Adaptation of respiratory patterns in collaborative reading. *In 14th Annual Conference of the International Speech Communication Association (Interspeech 2013)*, pages 1653–1657.

Bates, D., Machler, M., Bolker, B. et Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Benaglia, T., Chauveau, D., Hunter, D. R. et Young, D. (2009). mixtools : An R package for analyzing finite mixture models. *Journal of Statistical Software*, 32(6):1–29.

Benjamin, R. G. et Schwanenflugel, P. J. (2010). Text complexity and oral reading prosody in young readers. *Reading Research Quarterly*, 45(4):388–404.

Benjamin, R. G., Schwanenflugel, P. J., Meisinger, E. B., Groff, C., Kuhn, M. R. et Steiner, L. (2013). A spectrographically grounded scale for evaluating reading expressiveness. *Reading Research Quarterly*, 48(2):105–133.

Binder, K. S., Tighe, E., Jiang, Y., Kaftanski, K., Qi, C. et Ardoin, S. P. (2013). Reading expressively and understanding thoroughly : An examination of prosody in adults with low literacy skills. *Reading and writing*, 26(5):665–680.

Boersma, Paul & Weenink, D. (2020). Praat : doing phonetics by computer [computer program].

Bolanos, D., Cole, R. A., Ward, W. H., Tindal, G. A., Schwanenflugel, P. J. et Kuhn, M. R. (2013). Automatic assessment of expressive oral reading. *Speech Communication*, 55(2): 221–236.

Boliek, C. A., Hixon, T. J., Watson, P. J. et Jones, P. B. (2009). Refinement of speech breathing in healthy 4-to 6-year-old children. *Journal of Speech, Language, and Hearing Research*.

Bolinger, D. et Bolinger, D. L. M. (1989). *Intonation and Its Uses : Melody in Grammar and Discourse*. Stanford University Press.

Breen, M., Kaswer, L., Van Dyke, J. A., Krivokapić, J. et Landi, N. (2016). Imitated prosodic fluency predicts reading comprehension ability in good and poor high school readers. *Frontiers in psychology*, 7:1026.

Brennan, S. E. et Williams, M. (1995). The feeling of anothers knowing : Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of memory and language*, 34(3):383–398.

Bressoux, P., Bianco, M., Bosse, M., Cosnefroy, O., Dessus, P., Fayol, M., Hanner, C., Joët, G., Leroy, N., Lima, L., Massonnié, J. et Rocher, T. (2016). Rapport de la recherche longit.

Breznitz, Z. (1990). Vocalization and pauses in fast-paced reading. *The Journal of general psychology*, 117(2):153–159.

Calet, N., Gutiérrez-Palma, N. et Defior, S. (2017). Effects of fluency training on reading competence in primary school children : The role of prosody. *Learning and Instruction*.

Calet, N., Gutiérrez-Palma, N., Simpson, I. C., González-Trujillo, M. C. et Defior, S. (2015a). Suprasegmental phonology development and reading acquisition : A longitudinal study. *Scientific Studies of Reading*, 19(1):51–71.

Calet, N., Gutiérrez-Palma, N. et Defior, S. (2015b). A cross-sectional study of fluency and reading comprehension in spanish primary school children : Fluency and comprehension : Cross-sectional study. *Journal of Research in Reading*, 38:272–285.

Campione, E. et Véronis, J. (2002). A large-scale multilingual study of silent pause duration. *In Speech prosody*, pages 199–202.

Campione, E., Véronis, J. et DELIC, E. (2002). Étude des relations entre pauses et ponctuations pour la synthèse de la parole à partir de texte. *In Traitement Automatique des Langues Naturelles (TALN)*.

Chafe, W. (1988). Punctuation and the prosody of written language. *Written communication*, 5(4):395–426.

Chard, D. J., Vaughn, S. et Tyler, B.-J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of learning disabilities*, 35(5):386–406.

Cowie, R., Douglas-Cowie, E. et Wichmann, A. (2002). Prosodic characteristics of skilled reading : Fluency and expressiveness in 8—10-year-old readers. *Language and Speech*, 45(1):47–82.

Crunelle, D., Taillant, A. et Tiberghien, C. (2000). *Sacré Nestor*. OrthoEditions.

Cuetos, F., Martínez-García, C. et Suárez-Coalla, P. (2018). Prosodic perception problems in spanish dyslexia. *Scientific Studies of Reading*, 22(1):41–54.

de Boysson-Bardies, B. (1996). *Comment la parole vient aux enfants*. Odile Jacob.

de Boysson-Bardies, B., Sagart, L. et Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of child language*, 11(1):1–15.

Dellwo, V., Leemann, A. et Kolly, M.-J. (2015). The recognition of read and spontaneous speech in local vernacular : The case of zurich german. *Journal of Phonetics*, 48:13–28.

Dickens, R. H. et Meisinger, E. B. (2016). Examining the effects of skill level and reading modality on reading comprehension. *Reading Psychology*, 37(2):318–337.

Dodane, C. et Brunellière, A. (2006). Lecture silencieuse et oralisée des phrases relatives : le rôle de la prosodie. *Journées d'études sur la parole (JEP)*.

Dowd, A. J. et Bartlett, L. (2019). The need for speed : Interrogating the dominance of oral reading fluency in international reading efforts. *Comparative Education Review*, 63(2):189–212.

Dowhower, S. L. (1991). Speaking of prosody : Fluency's unattended bedfellow. *Theory into practice*, 30(3):165–175.

Dupoux, E., Peperkamp, S. et Sebastián-Gallés, N. (2001). A robust method to study stress "deafness". *The Journal of the Acoustical Society of America*, 110(3):1606–1618.

Encrevé, P. (1988). La liaison sans et avec enchainement. *Phonologie tridimensionnelle et usages du français*.

Erekson, J. A. (2010). Prosody and interpretation. *Reading Horizons*, 50(2):3.

FayoL, M. (1997). On acquiring and using punctuation : A study of written french. *Processing interclausal relationships. Studies in the production and comprehension of text*, pages 157–178.

Fernandes, S., Querido, L., Verhaeghe, A. et Araújo, L. (2018). What is the relationship between reading prosody and reading comprehension in european portuguese ? evidence from grades 2 to 5. *Journal of Research in Reading*.

Fernandes, S., Querido, L., Verhaeghe, A., Marques, C. et Araújo, L. (2017). Reading development in european portuguese : relationships between oral reading fluency, vocabulary and reading comprehension. *Reading and Writing*, 30(9):1987–2007.

Filipe, M. G., Peppé, S., Frota, S. et Vicente, S. G. (2017). Prosodic development in European Portuguese from childhood to adulthood. *Applied Psycholinguistics*, 38(05):1045–1070.

Fournier, D. A., Skaug, H., Ancheta, J., Ianelli, J., Magnussen, A., Maunder, M., Nielsen, A. et Sibert, J. (2012). Ad model builder : using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw.*, 27:233–249.

Frazier, L., Carlson, K. et Cliftonjr, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, 10(6):244–249.

Friot, B. (2007). *Histoires pressées*. Milan.

Fuchs, L. S., Fuchs, D., Hosp, M. K. et Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence : A theoretical, empirical, and historical analysis. *Scientific studies of reading*, 5(3):239–256.

Gentaz, E., Sprenger-Charolles, L., Theurel, A. et Colé, P. (2013). Reading comprehension in a large cohort of french first graders from low socio-economic status families : A 7-month longitudinal study. *PLoS One*, 8(11):e78608.

Godde, E., Bailly, G., Escudero, D., Bosse, M.-L., Bianco, M. et Vilain, C. (2017a). Improving fluency of young readers : introducing a karaoke to learn how to breath during a reading-while-listening task. *In Speech and Langage Technology for Education 2017*.

Godde, E., Bailly, G., Escudero, D., Bosse, M.-L. et Gillet-Perret, E. (2017b). Evaluation of reading performance of primary school children : Objective measurements vs. subjective ratings. *In International Workshop on Child Computer Interaction (WOCCI)*.

Godde, E., Bosse, M.-L. et Bailly, G. (2020a). A review of reading prosody acquisition and development. *Reading and Writing*, 33(2):399–426.

Godde, E., Bosse, M.-L. et Bailly, G. (2020b). Échelle multi-dimensionnelle de fluence : nouvel outil d'évaluation de la fluence en lecture prenant en compte la prosodie, étalonné du ce1 à la 5ème. *In press*.

Golan, O., Baron-Cohen, S. et Hill, J. (2006). The cambridge mindreading (cam) face-voice battery : Testing complex emotion recognition in adults with and without asperger syndrome. *Journal of autism and developmental disorders*, 36(2):169–183.

Goldman, J.-P., Auchlin, A. et Simon, A. C. (2009). Discrimination de styles de parole par analyse prosodique semi-automatique. *Actes d'IDP*, pages 207–221.

Goldman, J.-P., François, T., Roekhaut, S. et Simon, A. C. (2010). Étude statistique de la durée pausale dans différents styles de parole. *Journées d'Etude sur la Parole (JEP)*, pages 161–164.

Goswami, U., Fosker, T., Huss, M., Mead, N. et Szűcs, D. (2011). Rise time and formant transition duration in the discrimination of speech sounds : the ba–wa distinction in developmental dyslexia. *Developmental science*, 14(1):34–43.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S. et Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia : A new hypothesis. *Proceedings of the National Academy of Sciences*, 99(16):10911–10916.

Groen, M. A., Veenendaal, N. J. et Verhoeven, L. (2019). The role of prosody in reading comprehension : Evidence from poor comprehenders. *Journal of Research in Reading*, 42(1):37–57.

Grosjean, F. et Collins, M. (1979). Breathing, pausing and reading. *Phonetica*, 36(2):98–114.

Grosjean, F. et Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica*, 31(3-4):144–184.

Grosman, I., Simon, A. C. et Degand, L. (2018). Variation de la durée des pauses silencieuses : impact de la syntaxe, du style de parole et des disfluences. *Langages*, 211(3):13–40.

Guaïtella, I. (1999). Rhythm in speech : What rhythmic organizations reveal about cognitive processes in spontaneous speech production versus reading aloud. *Journal of pragmatics*, 31(4): 509–523.

Harris, M. (2013). *Language experience and early language development : From input to uptake*. Psychology Press.

Haskins, T. et Aleccia, V. (2014). Toward a reliable measure of prosody : an investigation of rater consistency. *Int J Educ Soc Sci*, 1(5):102–12.

Hayes, A. F. et Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1):77–89.

Hirschberg, J. (2000). A corpus-based approach to the study of speaking style. *In Prosody : Theory and experiment*, pages 335–350. Springer.

Hirst, D. et Di Cristo, A. (1998). *Intonation systems : a survey of twenty languages*. Cambridge University Press.

Hirst, D., Rilliard, A. et Aubergé, V. (1998). Comparison of subjective evaluation and an objective evaluation metric for prosody in text-to-speech synthesis. *In The Third ESCA/COCOSDA Workshop (ETRW) on Speech Synthesis*.

Holliman, A. J. (2014). Speech rhythm sensitivity in pre-readers : What role does it have in reading acquisition ?

Holliman, A. J., Palma, N. G., Critten, S., Wood, C., Cunnane, H. et Pillinger, C. (2017). Examining the independent contribution of prosodic sensitivity to word reading and spelling in early readers. *Reading and Writing*, 30(3):509–521.

Hoover, W. A. et Gough, P. B. (1990). The simple view of reading. *Reading and writing*, 2(2):127–160.

Hosp, M. K., Hosp, J. L. et Howell, K. W. (2016). *The ABCs of CBM : A practical guide to curriculum-based measurement*. Guilford Publications.

Howell, P. et Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication*, 10(2):163–169.

Jordán, N., Cuetos, F. et Suárez-Coalla, P. (2019). Prosody in the reading of children with specific language impairment/la prosodia en la lectura de niños con trastorno específico del lenguaje. *Infancia y Aprendizaje*, 42(1):87–127.

Kentner, G. et Vasishth, S. (2016). Prosodic Focus Marking in Silent Reading : Effects of Discourse Context and Rhythm. *Frontiers in Psychology*, 7.

Klauda, S. L. et Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational psychology*, 100(2):310.

Kocaarslan, M. (2019). The effects of reading rate, accuracy and prosody on second grade students. *Acta psychologica*, 197:86–93.

Koriat, A., Kreiner, H. et Greenberg, S. N. (2002). The extraction of structure during reading : Evidence from reading prosody. *Memory & cognition*, 30(2):270–280.

Kuhn, M. R. (2005). A comparative study of small group fluency instruction. *Journal of Investigative Surgery*, 26(2):127–146.

Kuhn, M. R. et Schwanenflugel, P. J. (2019). Prosody, pacing, and situational fluency (or why fluency matters for older readers). *Journal of Adolescent & Adult Literacy*, 62(4):363–368.

Kuhn, M. R., Schwanenflugel, P. J. et Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency : Automaticity, prosody, and definitions of fluency. *Reading research quarterly*, 45(2):230–251.

Kuhn, M. R. et Stahl, S. A. (2003). Fluency : A review of developmental and remedial practices. *Journal of educational psychology*, 95(1):3.

Lai, S. A., George Benjamin, R., Schwanenflugel, P. J. et Kuhn, M. R. (2014). The Longitudinal Relationship Between Reading Fluency and Reading Comprehension Skills in Second-Grade Children. *Reading & Writing Quarterly*, 30:116–138.

Lalain, M., Espesser, R., Ghio, A., De Looze, C., Reis, C. et Mendonca-Alves, L. (2014). Prosodie et lecture : particularités temporelles et mélodiques de l'enfant dyslexique en lecture et en narration. *Revue de Laryngologie Otologie Rhinologie, Revue de Laryngologie*, 135(2).

Lalain, M., Legou, T., Fauth, C., Hirsch, F. et Didirkova, I. (2016). Que disent nos silences ? apport des données acoustiques, articulatoires et physiologiques pour l'étude des pauses silencieuses. *In JEP-TALN-RECITAL 2016*, volume 1, pages 563–570.

Lalain, M., Mendonça-Alves, L., Espesser, R., Ghio, A., De, C. et Looze, C. R. (2012). Lecture et prosodie chez l'enfant dyslexique, le cas des pauses. *In Journées d'Etudes sur la Parole*, pages 41–48.

Lefavrais, P. (2005). *L'Alouette-R.* Centre de psychologie appliquée.

Levasseur, V. M., Macaruso, P. et Shankweiler, D. (2008). Promoting gains in reading fluency : A comparison of three approaches. *Reading and Writing*, 21(3):205–230.

Long, M. (2014). 'i can read further and there's more meaning while i read' : An exploratory study investigating the impact of a rhythm-based music intervention on children's reading. *Research Studies in Music Education*, 36(1):107–124.

Lopes, J. A., Silva, M. M., Moniz, A. V., Spear-Swerling, L. et Zibulsky, J. (2015). Prosody growth and reading comprehension : a longitudinal study from 2nd through the end of 3rd grade. *Revista de Psicodidáctica/Journal of Psychodidactics*, 20(1):5–23.

Lété, B., Sprenger-Charolles, L. et Colé, P. (2004). Manulex : A grade-level lexical database from french elementary-school readers. *Behavior Research Methods, Instruments, & Computers*, 36:156–166.

Maeder, C., Roustit, J., Launay, L. et Touzin, M. (2018). Evaleo 6-15.

Marston, D. B. (1989). *A curriculum-based measurement approach to assessing academic performance : What it is and why do it.* Guilford Press.

Martin, P. (2005). Winpitch ltl, un logiciel multimédia d'enseignement de la prosodie. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, 8(2).

Martin, P. (2009). *Intonation du français.* Armand Colin.

Martin, P. (2011). Ponctuation et structure prosodique. *Langue française*, 172:99–114.

Meisinger, E. B., Bradley, B. A., Schwanenflugel, P. J., Kuhn, M. R. et Morris, R. D. (2009). Myth and reality of the word caller : The relation between teacher nominations and prevalence among elementary school children. *School Psychology Quarterly*, 24(3):147.

Miller, J. et Schwanenflugel, P. J. (2006). prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology*, 98:839–853.

Miller, J. et Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading research quarterly*, 43(4):336–354.

Molomer, S. D., Trausan-Matu, S., Dessus, P. et Bianco, M. (2015). Analyzing students pauses during reading and explaining a story. *In 2015 14th RoEduNet International Conference-Networking in Education and Research (RoEduNet NER)*, pages 90–93. IEEE.

Morgan, J. L. et Demuth, K. (2014). *Signal to syntax : Bootstrapping from speech to grammar in early acquisition.* Psychology Press.

Morrison, T. G. et Wilcox, B. (2020). Assessing expressive oral reading fluency. *Education Sciences*, 10(3):59.

Moser, G. P., Sudweeks, R. R., Morrison, T. G. et Wilcox, B. (2014). Reliability of ratings of children's expressive reading. *Reading Psychology*, 35(1):58–79.

Nakamura, M., Iwano, K. et Furui, S. (2008). Differences between acoustic characteristics of spontaneous and read speech and their effects on speech recognition performance. *Computer Speech & Language*, 22(2):171–184.

Nelson, D. G. K., Hirsh-Pasek, K., Jusczyk, P. W. et Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of child Language*, 16(1):55–68.

O'Connor, R. E. (2018). Reading fluency and students with reading disabilities : How fast is fast enough to promote reading comprehension ? *Journal of learning disabilities*, 51(2):124–136.

Paige, D. D., Rasinski, T., Magpuri-Lavell, T. et Smith, G. S. (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research*, 46(2):123–156.

Paige, D. D., Rasinski, T. V. et Magpuri-Lavell, T. (2012). Is fluent, expressive reading important for high school readers ? *Journal of Adolescent & Adult Literacy*, 56(1):67–76.

Paige, D. D., Rupley, W. H., Smith, G. S., Rasinski, T. V., Nichols, W. et Magpuri-Lavell, T. (2017). Is prosodic reading a strategy for comprehension ? *Journal for Educational Research Online/Journal für Bildungsforschung Online*, 9(2):245–275.

Parker, R., Hasbrouck, J. E. et Tindal, G. (1992). The maze as a classroom-based reading measure : Construction methods, reliability, and validity. *The Journal of Special Education*, 26(2):195–218.

Patel, R., Kember, H. et Natale, S. (2014). Feasibility of augmenting text with visual prosodic cues to enhance oral reading. *Speech Communication*, 65:109–118.

Pinnell, G. S. *et al.* (1995). *Listening to children read aloud : Data from NAEP's integrated reading performance record (IRPR) at grade 4.* ERIC.

Prieto, P. et Esteve-Gibert, N. (2018). *The Development of Prosody in First Language Acquisition*, volume 23. John Benjamins Publishing Company.

Prior, S. M., Fenwick, K. D., Saunders, K. S., Ouellette, R., O'Quinn, C. et Harvey, S. (2011). Comprehension after oral and silent reading : Does grade level matter ? *Literacy Research and Instruction*, 50(3):183–194.

Rasinski, T. (2006). Reading fluency instruction : Moving beyond accuracy, automaticity, and prosody. *The Reading Teacher*, 59(7):704–706.

Rasinski, T., Paige, D., Rupley, W. et Young, C. (2019). Reading fluency : From theory to proof of concept, and from science to art. In 69th Literacy Research Association Annual Conference.

Rasinski, T., Rikli, A. et Johnston, S. (2009). Reading fluency : More than automaticity ? more than a concern for the primary grades ? *Literacy Research and Instruction*, 48(4):350–361.

Rasinski, T. V. (2004). Assessing reading fluency. *Pacific Resources for Education and Learning (PREL).*

Rasinski, T. V. (2010). *The fluent reader : Oral & silent reading strategies for building fluency, word recognition & comprehension.* Scholastic.

Rasinski, T. V., Padak, N. D., McKeon, C. A., Wilfong, L. G., Friedauer, J. A. et Heim, P. (2005). Is reading fluency a key for successful high school reading ? *Journal of Adolescent & Adult Literacy*, 49(1):22–27.

Rasinski, T. V. et Smith, M. C. (2018). *The megabook of fluency : Strategies and texts to engage all readers.* Scholastic.

Raven, J. *et al.* (2003). Raven progressive matrices. *In Handbook of nonverbal assessment*, pages 223–237. Springer.

Ravid, D. et Mashraki, Y. E. (2007). Prosodic reading, reading comprehension and morphological skills in hebrew-speaking fourth graders. *Journal of Research in Reading*, 30(2):140–156.

Reilhac, C., Bosse, M. et Valdois, S. (2018). Independent contribution of visual attention span, phoneme awareness and rapid automatized naming deficits to poor reading. *submitted.*

Rossi, M. (1980). Le français, langue sans accent ? *Studia Phonetica Montréal*, 15:13–51.

Schimmel, N. et Ness, M. (2017). The effects of oral and silent reading on reading comprehension. *Reading Psychology*, 38(4):390–416.

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior*, 12(3):177–186.

Schreiber, P. A. (1987). *Prosody and structure in children's syntactic processing.* Academic Press.

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into practice*, 30(3):158–164.

Schwanenflugel, P., Brock, M., Tanaka, V., Westmoreland, M. et Mon, S. (2016). Influence of passage genre on reading prosody. *In Twenty-Third Annual Meeting Society for the Society for the Scientific Study of Reading, Porto, Portugal*.

Schwanenflugel, P. J. et Benjamin, R. G. (2012). Reading Experssiveness : The Neglected aspect of Reading Fluency. *In Fluency Instruction : Research Based Best Practice*. Guilford Press.

Schwanenflugel, P. J. et Benjamin, R. G. (2017). Lexical prosody as an aspect of oral reading fluency. *Reading and Writing*, 30(1):143–162.

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M. et Stahl, S. A. (2004). Becoming a Fluent Reader : Reading Skill and Prosodic Features in the Oral Reading of Young Readers. *Journal of Educational Psychology*, 96(1):119–129.

Schwanenflugel, P. J., Westmoreland, M. R. et Benjamin, R. G. (2015). Reading fluency skill and the prosodic marking of linguistic focus. *Reading and Writing*, 28:9–30.

Seymour, P. H., Aro, M. et Erskine, J. M. (2003). Foundation literacy acquisition in european orthographies. *British Journal of psychology*, 94(2):143–174.

Smith, G. S. et Paige, D. D. (2019). A study of reliability across multiple raters when using the naep and mdfs rubrics to measure oral reading fluency. *Reading Psychology*, 40(1):34–69.

Smithson, M. et Verkuilen, J. (2006). A better lemon squeezer ? maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods*, 11(1):54.

Suárez-Coalla, P., Álvarez-Cañizo, M., Martínez, C., García, N. et Cuetos, F. (2016). Reading prosody in spanish dyslexics. *Annals of dyslexia*, 66(3):275–300.

Team, R. (2019). R : A language and environment for statistical computing. vienna, austria : R foundation for statistical computing ; 2011.

Veenendaal, N. J., Groen, M. A. et Verhoeven, L. (2014). The role of speech prosody and text reading prosody in children's reading comprehension. *British Journal of Educational Psychology*, 84:521–536.

Veenendaal, N. J., Groen, M. A. et Verhoeven, L. (2016a). Bidirectional relations between text reading prosody and reading comprehension in the upper primary school grades : A longitudinal perspective. *Scientific Studies of Reading*, 20(3):189–202.

Veenendaal, N. J., Groen, M. A. et Verhoeven, L. (2016b). The contribution of segmental and suprasegmental phonology to reading comprehension. *Reading research quarterly*, 51(1):55–66.

Wang, Y.-T., Green, J. R., Nip, I. S., Kent, R. D. et Kent, J. F. (2010). Breath group analysis for reading and spontaneous speech in healthy adults. *Folia Phoniatrica et Logopaedica*, 62(6):297–302.

Wells, B., Peppé, S. et Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, 31:749–778.

Wennerstrom, A. (2001). *The music of everyday speech : Prosody and discourse analysis*. Oxford University Press.

Wiechern, B., Liberty, K. A., Pattemore, P. et Lin, E. (2018). Effects of asthma on breathing during reading aloud. *Speech, Language and Hearing*, 21(1):30–40.

Wlodarczak, M. et Heldner, M. (2017). Respiratory constraints in verbal and non-verbal communication. *Frontiers in psychology*, 8:708.

Wood, C., Wade-Woolley, L. et Holliman, A. J. (2009). Phonological awareness. *Contemporary perspectives on reading and spelling*, pages 7–23.

Yildirim, K., Rasinski, T. et Kaya, D. (2018). Fluency and comprehension of narrative texts in turkish students in grades 4 through 8. *Education 3-13*, pages 1–10.

Yildiz, M., Yildirim, K., Ates, S., Rasinski, T., Fitzgerald, S. et Zimmerman, B. (2014). The relationship between reading fluency and reading comprehension in fifth-grade turkish students. *International Journal of School & Educational Psychology*, 2(1):35–44.

Young, A. R. et Bowers, P. G. (1995). Individual Difference and text difficulty determinant of reading fluency and expressiveness. *Journal of Experimental Child Psychology*, 60(3):428–454.

Young, A. R., Bowers, P. G. et MacKinnon, G. E. (1996). Effects of prosodic modeling and repeated reading on poor readers' fluency and comprehension. *Applied Psycholinguistics*, 17:59–84.

Young, C. et Rasinski, T. (2018). Readers theatre : Effects on word recognition automaticity and reading prosody. *Journal of Research in Reading*, 41(3):475–485.

Young, C., Valadez, C. et Gandara, C. (2016). Using performance methods to enhance students' reading fluency. *The Journal of Educational Research*, 109(6):624–630.

Young-Suk, K., Quinn, J. M. et Petscher, Y. (2020). Reading prosody unpacked : A longitudinal investigation of its dimensionality and relation with word reading and listening comprehension for children in primary grades. *Journal of Educational Psychology*.

Zorman, M., Lequette, C., Pouget, G., Devaux, M. et Savin, H. (2008). Entraînement de la fluence de lecture pour les élèves de 6e en difficulté de lecture. *ANAE, 96*, 97:213–219.

Zutell, J. et Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice*, 30(3):211–217.

Zvonik, E. et Cummins, F. (2003). The effect of surrounding phrase lengths on pause duration. *In Eighth European conference on speech communication and technology*.

Annexe A

Textes utilisés pour les lectures prosodiques

Texte 1

Mon chat est entré dans ma chambre. Il avait l'air bizarre donc j'ai tout de suite compris qu'il avait fait une énorme bêtise. D'abord j'ai pensé au poisson rouge. Je m'attendais à trouver le bocal vide. Mais mon poisson était toujours là et il me regardait de ses grands yeux ronds. Ensuite je me suis souvenu que maman avait sorti des saucisses et les avait posées sur la table. Le chat les avait volées, c'était certain. Mais dans la cuisine, les saucisses étaient toujours là, prêtes à cuire, personne n'y avait touché. Le chat est alors allé vers mon cartable. Le cartable était ouvert, tout était par terre, mon exposé déchiré en tout petits bouts. Quand j'ai vu mon travail en morceaux, je me suis senti triste, j'avais envie de pleurer. Alors le chat s'est frotté sur ma jambe, il a miaulé, il était désolé. Finalement il m'a aidé et on a préparé ensemble un bel exposé sur l'intelligence des chats.

Texte 2

Lucas arrive à l'école. Il retrouve son amie Lola.
« - Salut Lola ! Est-ce que tu as préparé ton exposé ? demande Lucas.
- Oui, j'ai parlé de mon livre préféré, répond Lola. Et toi ? Je suis sûre que tu as choisi les super-héros !
- Et non ! Qui t'as dit que j'aimais les super héros ? s'étonne Lucas
- Enzo, je crois. Tu l'as fait sur quoi alors ton exposé ? poursuit Lola.
- Il se mêle toujours de tout celui-là, grogne Lucas.
- Allez, dis-moi ! Quel sujet as-tu choisi finalement ? reprend Lola. Les lions ? La danse ? La musique?
- C'est une histoire amusante, répond Lucas. Est-ce que je t'ai déjà parlé de mon chat ?
- C'est celui qui fait toujours des bêtises ? Je l'ai déjà vu chez toi, se souvient Lola.
- Oui, c'est bien lui. Hier il a mangé mon exposé sur les super-héros. Mais pour se faire pardonner, il m'a aidé à en faire un nouveau sur les chats ! »

Annexe B

Test de vocabulaire

Issu de l'étude Ortholearn (Reilhac *et al.*, 2018)

Nom : **Prénom :**

Coche la case du mot qui a le même sens que le mot du haut.

Exemples :

1. voiture

☐ camion ☐ coiffure ☐ automobile ☐ gardien ☐ véhicule

2. stupide

☐ absurde ☐ géant ☐ méchant ☐ premier ☐ idiot

Réponds le plus vite possible

Attends le signal pour tourner la page

1. animal

☐ **cannibale** ☐ **loup** ☐ **œuf** ☐ **bête** ☐ **dictée**

2. miroir

☐ défaite ☐ glace ☐ reflet ☐ mirage ☐rétroviseur

3. vase

☐ pot ☐ vache ☐ seau ☐ fleur ☐ plante

4. surprise

☐ étonnement ☐ cerise ☐ émotion ☐ chaleur ☐ stupeur

5. étang

☐ flaque ☐ citron ☐ mer ☐ lac ☐ étage

6. médicament

☐ docteur ☐ pommade ☐ médiane ☐ couleur ☐ remède

7. bâtiment

☐ aliment ☐ chouette ☐ ville ☐ cabane ☐ immeuble

8. fragile

☐ délicat ☐ cassé ☐ léger ☐ argile ☐ poilu

9. chaussure

☐ pied ☐ soulier ☐ chanson ☐ bruit ☐ pantoufle

10. flacon

☐ parfum ☐ café ☐ bouteille ☐ flocon ☐ verre

11. paquebot

☐ transport ☐ navire ☐ coquelicot ☐ barque ☐ paquet

12. jadis

☐ avant ☐ autrefois ☐ jamais ☐ déjà ☐ hélas

13. applaudir

☐ acclamer ☐ répondre ☐ accueillir ☐ aplanir ☐ saluer

14. souhaiter

☐ désirer ☐ penser ☐ demander ☐ soutenir ☐ prendre

15. alligator

☐ serpent ☐ papillote ☐ aspirateur ☐ crocodile ☐ animal

16. certain

☐ parfois ☐ vrai ☐ possible ☐ serein ☐ sûr

17. surveiller

☐ voir ☐ suivre ☐ placer ☐ contrôler ☐ survenir

18. potage

☐ partage ☐ légume ☐ soupe ☐ bougie ☐ dîner

19. dispute

☐ violence ☐ marin ☐ querelle ☐ distance ☐ combat

20. matière

☐ cours ☐ espace ☐ maths ☐ discipline ☐ matinée

21. divan

☐ divers ☐ canapé ☐ meuble ☐ lit ☐ fleuve

22. mobile

☐ déplaçable ☐ libre ☐ pratique ☐ solide ☐ malade

23. sondage

☐ série ☐ montage ☐ enquête ☐ opinion ☐ question

24. âtre

☐ buche ☐ cheminée ☐ plâtre ☐ feu ☐ maillot

25. tyran

☐ dictateur ☐ monstre ☐ tirage ☐ pirate ☐ roi

26. courge

☐ légume ☐ chou ☐ tapis ☐ courbe ☐ potiron

27. solliciter

☐ dire ☐ vouloir ☐ éclater ☐ expliquer ☐ demander

28. ajourner

☐ remettre ☐ juger ☐ révéler ☐ reculer ☐ séjourner

29. visionner

☐ filmer ☐ espionner ☐ regarder ☐ couper ☐enregistrer

30. idyllique

☐ agréable ☐ éthylique ☐ paradisiaque ☐ majestueux ☐ parfait

Annexe C

Tests de compréhension écrite globale

Issus de l'étude Longit (Bressoux *et al.*, 2016)

Le pêcheur

Anne et Jean cherchaient des têtards. Soudain, ils ont entendu un grand plouf. Un pêcheur était tombé dans le lac. Il ne pouvait pas nager car il s'était blessé. Les enfants ont essayé de le tirer vers le bord mais il était trop lourd. Alors, Anne a tenu la tête de l'homme hors de l'eau pendant que Jean a couru chercher de l'aide.

Nom : Prénom :

1- Que faisaient Jean et Anne près du lac ?

……

2- Quel bruit ont-ils entendu ?

……

3- Pourquoi Jean et Anne ont-ils entendu un grand plouf ?

……

4- Pourquoi le pêcheur ne peut-il pas nager ?

……

5- Qu'essayent de faire les enfants ?

……

6- Pourquoi sont-ils incapable de tirer le pêcheur vers le bord ?

……

7- Comment Anne aide-t-elle le pêcheur ?

……

8- Comment Jean l'aide-t-il ?

……

Au parc, Simon et Tania ont vu un homme qui courait. Un morceau de papier est tombé de sa poche. Les enfants l'ont ramassé et ont lu un mystérieux message : *Un trésor qui a du nez. Les pieds du bouddha.*

- « C'est sûrement un message codé, dit Simon. Il doit indiquer où a été caché un trésor. »

Les enfants, qui ne savaient pas ce qu'était un bouddha, allèrent demander à Marion, la grande sœur de Simon.

- Bouddha était un homme très bon qui savait beaucoup de choses, dit-elle. Dans la carrière de sable de Fontainebleau, il y a une grande statue de Bouddha. Si vous voulez, je vais vous la montrer.

Ils prirent tous les trois leur vélo. Arrivés à la carrière, ils se retrouvèrent à l'entrée d'un étrange couloir à ciel ouvert, entre deux falaises. Le couloir était très sombre. Simon et Tania se rapprochèrent l'un de l'autre pour se donner du courage et entrèrent. Soudain, un hurlement jaillit du fond du couloir.

- On dirait un cri d'animal, dit Marion. Peut-être un renard prisonnier d'un piège ?
- Il faut aller voir, affirma Simon, d'une voix qui tremblait un peu.

Ils avancèrent encore un peu jusqu'à se retrouver face à la statue de Bouddha. Le hurlement continuait.

Tania se baissa et dans l'obscurité, elle distingua, sous la statue, une cage, puis deux yeux qui la suppliaient et … le museau d'un petit chien !

- Oh, le pauvre, dit-elle, il faut le sortir de là !

Simon, Tania et Marion attrapèrent les barreaux et tirèrent de toutes leurs forces pour sortir la cage. Puis Tania ouvrit la porte. Le chien bondit avec une force imprévue et courut à toute vitesse.

Les enfants le retrouvèrent près de leurs vélos en train de croquer les biscuits de leur goûter.

- Il devait garder le trésor, dit Tania.
- Non, dit Simon. Dans la cachette, il n'y avait qu'une cage. Le trésor, c'est forcément lui !
- Il n'a pas l'air d'avoir une grande valeur, il est tout petit », dit Tania, déçue. Marion montra les restes du goûter.
- Je ne sais pas si c'est un chien de pure race mais il est allé tout droit sur mes gâteaux préférés. Quel flair ! Quel nez !

Simon dansa d'excitation :

- Marion a raison, le message disait : *Un trésor qui a du nez*

1

NOM : **Prénom :**

1) **Quand Simon dit « il doit indiquer où a été caché un trésor », ça veut dire que …**

- Simon doit indiquer où est caché un trésor ☐[a]
- Le message doit indiquer où est caché un trésor ☐[b]
- L'homme doit indiquer où est caché un trésor ☐[c]
- Le bouddha doit indiquer où est caché un trésor ☐[d]

2***)*** **Pourquoi Simon et Tania vont-ils voir Marion ?**

- Pour lui demander ce que veut dire le message ☐[a]
- Pour lui demander quel est le trésor ☐[b]
- Pour lui demander ce qu'est un bouddha ☐[c]
- Pour lui demander de garder le papier ☐[d]

3) **Pourquoi les enfants vont-ils à la carrière de Fontainebleau ?**

……
……
……

4) **Comment les enfants se rendent-ils à la carrière ?**

……
……
……

5) **Simon a la voix qui tremble un peu parce que …**

- Il a froid ☐[a]
- Il a peur ☐[b]
- Il est fatigué ☐[c]
- Il est essoufflé ☐[d]

1

***6)* Dans la carrière, qui hurlait ?**

- Le renard ☐[a]
- Simon ☐[b]
- Le chien ☐[c]
- Bouddha ☐[d]

***7)* Que trouvent les enfants sous la statue ?**

……
……
……

***8)* Quel trésor les enfants ont-ils trouvé ?**

……
……
……

***9)* Finalement, pourquoi le message disait « Un trésor qui a du nez. Les pieds du bouddha » ?**

- Le trésor est un chien avec un grand museau ; il est caché sous la statue du Bouddha ☐[a]
- Un chien qui a du flair vous conduira au trésor du Bouddha ☐[b]
- Le nez et les pieds de la statue du Bouddha cachent un trésor ☐[c]
- Le trésor est un chien qui a du flair ; il est caché sous la statue du Bouddha ☐[d]

Nous sommes en 2096. Depuis 2022, l'humanité s'est réfugiée dans un monde souterrain et c'était la onzième fois que les membres de l'ARES, l'*Association pour la Remontée En Surface*, défilaient devant le collège du Monde Souterrain où étudiait Élodie. Les manifestants se battaient pour que l'on remonte vivre sur Terre. Des scientifiques de l'association s'étaient, paraît-il, rendus dans le Monde d'En Haut pour y effectuer des mesures. Ils assuraient que les Grandes Pollutions qui avaient ravagé la Terre en 2022 en causant des millions de morts étaient presque toutes résorbées et qu'il était maintenant possible d'y revivre. On les avait d'abord pris pour de doux rêveurs. Mais peu à peu, l'idée de remonter vivre sur Terre avait fait son chemin et l'ARES avait regroupé de plus en plus de sympathisants.

Le gouvernement de Suburba avait alors publié plusieurs communiqués en affirmant que toutes les études sérieuses montraient que la Terre ne serait pas habitable avant plusieurs siècles. Les principaux membres de l'association avaient été emprisonnés et les soudures des énormes portes blindées qui donnaient accès au Monde d'En Haut avaient été renforcées.

Un garde qui surveillait la manifestation s'approcha d'Élodie. Il renversa son cartable d'un geste brusque et feuilleta rapidement ses cahiers. Les dents serrées, Élodie replaçait ses affaires dans son cartable au fur et à mesure que le garde les examinait. Il termina par un petit portefeuille de tissu dont Élodie ne se séparait jamais.

- « Qu'est-ce que c'est que ça ? », demanda-t-il en sortant une photo qu'il lui mit sous le nez.

- « Ça ? … C'est la maison de mon arrière-grand-père. À l'époque où il habitait le Monde d'En Haut. »

Élodie tenait beaucoup à cette photo. La maison de son arrière-grand-père semblait tout droit sortie d'un conte, petite et pleine de trucs incroyablement anciens dont elle ne connaissait même pas le nom. Elle avait toujours pensé qu'on devait s'y sentir bien. Dad, son grand-père, lui avait donné la photo quelques mois avant sa mort.

Le garde s'approcha de son chef, la photo à la main. Ils échangèrent quelques mots puis l'homme revient vers elle.

« Tu sais très bien que ces photos sont interdites, aboya-t-il, les seules photos du Monde d'En Haut autorisées sont celles des musées. Tes parents pourraient être condamnés à une très lourde amende à cause de ça ! »

Élodie hocha la tête. Totalement impuissante, elle regarda l'homme déchirer la photo en petits morceaux qu'il jeta à la poubelle.

Nom : Prénom :

1) **Pour quelle raison les hommes vivent-ils sous terre ?**

- À cause de l'ARES ☐[a]
- À cause de la guerre ☐[b]
- Pour faire des expériences scientifiques ☐[c]
- À cause de la pollution ☐[d]

2) **Que veulent les manifestants ?**

……

……

……

3) **Comment s'appelle le pays du Monde Souterrain ?**

……

……

……

4) **Pourquoi les soudures des portes blindées du Monde Souterrain ont-elles été renforcées ?**

- Parce qu'elles étaient usées ☐[a]
- Pour que les idées de l'ARES ne soient pas diffusées ☐[b]
- Pour que les prisonniers ne puissent pas s'échapper ☐[c]
- Pour que personne ne rejoigne le Monde d'En Haut ☐[d]

5) **Où le garde trouve-t-il la photo d'Élodie ?**

……

……

……

6) **Qui a habité dans le Monde d'En Haut ?**

……

……

……

7) **Pourquoi Élodie serre-t-elle les dents quand le garde fouille son cartable ?**

- Elle a peur que ses affaires soient abîmées ☐[a]
- Elle a mal aux dents ☐[b]
- Elle ne veut pas manifester ☐[c]
- Elle a peur que le garde trouve sa photo ☐[d]

8) **Que risquent les parents d'Élodie si elle garde sa photo ?**

……

……

……

9) **Pourquoi le garde déchire-t-il la photo d'Élodie ?**

……

……

……

10) **Pourquoi les principaux membres de l'ARES ont-ils été emprisonnés ?**

- Ils ont manifesté ☐[a]
- Ils disent qu'il est possible de revivre dans le Monde d'en Haut ☐[b]
- Ils ont attaqué le collège ☐[c]
- Ils disent que le Monde d'en Haut est trop pollué ☐[d]

Le navire du jeune capitaine Dick Sand avait fait naufrage sur les côtes africaines. Chaque jour, il explorait les alentours à la recherche de nourriture. Dans la journée du 9 juillet, il eut à faire preuve d'un extrême sang-froid. Dick Sand était seul à terre, à l'affût d'un caama dont les cornes se montraient au-dessus d'un taillis, et il venait de le tirer, lorsque bondit, à trente pas, un formidable chasseur, qui sans doute venait réclamer sa proie et n'était pas d'humeur à l'abandonner. C'était un lion de grande taille, de ceux que les indigènes appellent « karamos », et non de cette espèce sans crinière, dite « lion du Nyassi ». Celui-là mesurait cinq pieds de haut, une bête extraordinaire.

Du bond qu'il avait fait, le lion était tombé sur le caama que la balle de Dick Sand venait de jeter à terre, et qui, plein de vie encore, palpitait en criant sous la patte du terrible animal. Dick Sand, désarmé, n'avait pas eu le temps de glisser une seconde cartouche dans son fusil. Du premier coup, le lion l'avait aperçu, mais il se contenta d'abord de le regarder. Dick Sand fut assez maître de lui pour ne pas faire un mouvement. Il se souvint qu'en pareille circonstance l'immobilité peut être le salut. Il ne tenta pas de recharger son arme, il n'essaya même pas de fuir.

Le lion le regardait toujours de ses yeux de chat, rouges et lumineux. Il hésitait entre deux proies, celle qui remuait et celle qui ne remuait pas. Si le caama ne se fût pas tordu sous la griffe du lion, Dick Sand eût été perdu. Deux minutes s'écoulèrent ainsi. Le lion regardait Dick Sand, et Dick Sand regardait le lion, sans même remuer ses paupières. Et alors, d'un superbe coup de gueule, le lion, enlevant le caama tout pantelant, l'emporta comme un chien eût fait d'un lièvre, et, battant les arbustes de sa formidable queue, il disparut dans le haut taillis.

1

Nom : **Prénom :**

1) **Qui est Dick Sand ?**

……

……

……

2) **Sur quoi Dick Sand tire-t-il ?**

- Un lièvre ☐[a]
- Un lion du Nyassi ☐[b]
- Un karamos ☐[c]
- Un animal à cornes ☐[d]

3) **Au début du texte, qui est le formidable chasseur ?**

……

……

……

4) **Comment les indigènes appellent-ils le lion de grande taille ?**

……

……

……

5) **Pourquoi Dick Sand ne tire-t-il pas sur le lion ?**

……

……

……

6) **Quelle est la proie finalement choisie par le lion ?**

……

……

……

2

7) **Comment Dick Sand fait-il preuve de sang-froid ?**

……

……

……

8) **Dick Sand a-t-il réussi à ramener de la nourriture ?**

……

……

……

9) **À la fin de l'histoire, que fait le lion du caama ?**

……

……

……

10) **À ton avis, que ressent Dick Sand pendant que le lion le regarde ?**

- Il est calme ☐[a]
- Il est admiratif ☐[b]
- Il est effrayé ☐[c]
- Il est soulagé ☐[d]

Texte 7 (CE2 et CM1)

Un jour, dans une famille où il y avait déjà sept garçons, naquit une petite fille. L'enfant était si fragile et si menue que ses parents décidèrent de lui faire boire, pour la fortifier, de l'eau d'une fontaine magique.

Ils y envoyèrent leurs fils remplir une cruche. En chemin, les garçons décidèrent de faire une course et ce fut une belle bousculade ! Hélas ! Ils firent tomber la cruche et la brisèrent en mille morceaux. Les pauvres enfants, épouvantés à l'idée d'affronter la colère de leur père, restaient pétrifiés de crainte. Lui, de son côté, ne les voyant pas rentrer, s'impatientait.

- « Ces sacrés garnements ! disait-il. On ne peut jamais compter sur eux ! Ils n'ont pas plus de cervelle que des oiseaux ! Que ne sont-ils pas plutôt des corbeaux ! ». Au moment même, il vit sept corbeaux s'éloigner dans le ciel. C'était ses fils que ses paroles imprudentes avaient changés en oiseaux. Les parents en conçurent un chagrin extrême.

Des années plus tard, la mère raconta à sa fille ce qui s'était passé. L'enfant décida de retrouver ses frères. Elle partit un matin, emportant une bague aux vertus singulières que sa mère lui avait donnée. Elle marcha longtemps et arriva chez l'Étoile du Matin.
- « Tes frères sont chez le nain de la montagne, lui dit l'Étoile. Si tu parviens au sommet où il habite, montre ton anneau à tes frères. S'ils le reconnaissent, le charme qui les retient prisonniers sera brisé ».

La fillette escalada la montagne. Lorsqu'elle arriva au sommet, elle frappa à la porte. Le nain lui ouvrit et la fit entrer. Sept écuelles et sept gobelets étaient posés sur la table. La petite fille cacha l'anneau dans l'un des gobelets et se dissimula dans un coin de la pièce. Les oiseaux arrivèrent bientôt et commencèrent à manger et à boire. Soudain, l'un deux s'écria : « Regardez le bel anneau que j'ai trouvé ! ». Alors, le plus âgé des corbeaux le prit dans son bec et dit : « c'est l'anneau de notre mère, Dieu soit loué ! ».

À l'instant même, les sept frères retrouvèrent leur figure humaine. Leur sœur se précipita pour les embrasser tous. Après avoir remercié le nain pour ses bontés, ils repartirent joyeusement vers la maison de leurs parents qui furent bien aise de revoir leurs enfants.

Nom : …….. Prénom : ……..

1) **Au début de l'histoire, où les parents envoient-ils leurs fils ?**

……

……

……..

2) **Que font les garçons avec la cruche ?**

……

……

……..

3) **Au début de l'histoire, pourquoi le père s'impatiente-t-il ?**

- Il attend que ses garçons rentrent . . . . . □[a]
- Il attend ses garçons pour aller faire une course avec eux . . . . . □[b]
- Il attend ses garçons pour partir remplir une cruche . . . . . □[c]
- Il attend que sa femme lui serve à boire . . . . . □[d]

4) **Qui n'a pas plus de cervelle que des oiseaux ?**

……

……

……..

5) **Les parents sont tristes parce que …**

- Les parents ont été transformés en corbeaux . . . . . □[a]
- Les corbeaux ont cassé la cruche. . . . . □[b]
- Leurs fils ne savent pas compter. . . . . □[c]
- Leurs enfants ont été transformés en corbeaux. . . . . □[d]

6) **Que fait la petite fille avec l'anneau ?**

……

……

……..

7) **Comment les sept frères retrouvent-ils leur figure humaine ?**

- Le nain a prononcé une formule magique. . . . . □[a]
- Les corbeaux ont prononcé une formule magique . . . . . □[b]
- Ils ont bu de la potion magique dans leur écuelle . . . . . □[c]
- Ils ont reconnu l'anneau de leur mère. . . . . □[d]

8) **Où les sept frères ont-ils vécu pendant des années ?**

……

……

……..

9) **Qui retrouve les sept frères ?**

……

……

……..

10) **Qui est responsable de la transformation des frères en corbeaux ?**

……

……

……..

Texte 2

Un jour, à Londres, un homme se prépare à faire un voyage en train tandis qu'une jeune femme se rend dans un magasin.

La première cliente du magasin, une jeune femme qui était secrétaire dans le centre-ville, choisit huit oranges bien dodues. Après tout, aujourd'hui c'était son anniversaire. Après le travail, elle fut invitée par ses collègues à boire un verre pour fêter son anniversaire. Le sac de fruits l'accompagna. Elle se promit de ne rester qu'un moment mais une heure et trois verres de vin plus tard, elle était encore au bar. Elle n'était pas habituée à boire autant d'alcool. Sa tête tournait. Quand elle réalisa qu'il était temps de partir, elle faillit laisser ses oranges. Lorsqu'elle prit le chemin de la gare, elle ne tenait plus très bien sur ses jambes.

John serrait fermement son colis contre lui et se fraya un passage vers l'autre bout du quai, là où il y avait moins de voyageurs. John avait estimé qu'avec un objet si précieux, il devait se rendre en personne à la vente aux enchères.

Sitôt arrivée, la secrétaire vit son train qui approchait. Au même moment, John s'avança vers la ligne blanche qui marquait le bord du quai. Il y eut un mouvement général vers l'avant. Dans la foule, la secrétaire, encore troublée par le vin, fut bousculée par des voyageurs. Le sac en papier se déchira et les fruits s'éparpillèrent.

Une des oranges roula vers John qui étreignait son précieux bagage. L'orange arriva au moment précis où il s'avançait encore un peu plus près du bord. Son pied glissa dessus. Il perdit l'équilibre et s'efforça de ne pas lâcher son colis. Surtout ne pas lâcher le colis ! Pendant une demi-seconde, John oscilla sur le bord du quai. Puis il dégringola sur les rails. Le conducteur freina de toutes ses forces mais le train allait trop vite. Ce fut fini en un instant. Des cris d'effroi retentirent. Blanche comme un linge, la secrétaire plaqua sa main sur sa bouche grande ouverte. L'orange était intacte.

Nom : ..... Prénom : .....

Réponds aux questions sur les pages suivantes

1) Pourquoi la secrétaire va-t-elle boire un verre avec ses collègues ?

_____

_____

_____

2) Pourquoi la secrétaire ne tient plus très bien sur ses jambes ?

- Elle ne connaît pas le chemin de la gare ☐ a
- Elle a bu trop de vin ☐ b
- Elle a la tête qui tourne ☐ c
- Elle est trop chargée ☐ d

3) Que contient le colis de John ?

_____

_____

_____

4) Pourquoi John prend le train avec son colis ?

_____

_____

_____

5) Quels personnages s'apprêtent à prendre le train ?

_____

_____

_____

6) Pourquoi John perd-il l'équilibre ?

_____

_____

_____

7) À la fin de l'histoire, qu'arrive-t-il à John ?

_____

_____

_____

8) La secrétaire est blanche comme un linge parce que …

- Elle a perdu toutes ses oranges ☐ a
- Elle a vu John marcher sur ses oranges ☐ b
- Elle a vu John passer sous le train ☐ c
- Elle a bu trop d'alcool ☐ d

9) Qui est responsable de l'accident de John ?

_____

_____

_____

10) John et la secrétaire se connaissent-ils ?

_____

_____

_____

Annexe D

Maze-tests pour la compréhension écrite locale

Nom : …….. Prénom :……..

Consigne : Entoure le mot qui convient.

Exemple :

Il aime manger des **voitures**. Sa maman lui en **bonbons** / **caramel** / **donne** / **grossir** / **voit** souvent .

C’est vraiment ennuyeux de se **cacher** / **pause** / **lever** le matin. Comme d’habitude, je **lui** / **suis** / **peux** descendu prendre mon petit déjeuner, mon **bol** / **assis** / **colle** de chocolat m’attendait sur la **classe** / **table**. Ensuite, maman m’a dit de **le** / **mon** / **me** dépêcher de me laver parce que **filons** / **sinon** / **non** j’allais être en retard à l’ **inviter** / **arrivée** / **école**. Elle me dit la même chose **sous** / **nous** / **tous** les jours, mais rien ne me **regarde** / **retarde** / **bavard** jamais sur le chemin de l’école.

Fait **la**
Mais aujourd'hui, quand je suis descendu dans **ça** rue, la voisine d'en face était
Vers **lit**

dessous **vite**
ressort, en robe de chambre, paniquée. J'ai **suite** compris que son minet
dehors **hier**

gagné **de**
adoré avait **grimpé** dans l'arbre le plus grand **me** la rue et ne savait plus
grain **à**

faire **escalader**
quoi descendre. Moi, je sais grimper aux **rochers** , alors je suis allé le
comment **arbres**

En **corbeau** **rond**
chercher. **Quand** je suis arrivé à côté du **monter**, il a sauté sur mon dos, **lui** est
Qui **chat** **on**

dû **on**
descendu comme ça. Ensuite, j'ai **bu** courir pour aller à l'école, et **je** me suis
pull **lui**

chauffeur **commencée**
fais gronder par le **faché** parce que la dictée était déjà **attrapée** quand je
maitre **cahier**

Car **par**
suis entré en classe. **Vrai** ce n'était pas très grave : **pour** une fois, quelque
Mais **prêter**

arrivé **mentais**
chose d'exceptionnel était **content** dans ma journée, et je me **sentais** l'âme
parti **forêt**

d'un héros.

Nom : …….. Prénom :……..

Consigne : Entoure le mot qui convient.

Exemple :

bonbons **caramel**

Il aime manger des **voitures**. Sa maman lui en **donne** souvent .

grossir **voit**

---

habité

Un homme cagoulé s'est introduit dans l' **ascenseur** de Rosa Tavu, la célèbre

appartement

Elle **pétanque**

voyante. **Il** lui a dérobé sa boule de **demain** Lilavenir, dernier modèle, qui

Pour **cristal**

prédire **moquée**

permet de **sauver** l'avenir avec une fiabilité inégalée. Jalousée **avec** toutes

futur **par**

cette

ses collègues parce qu'elle possédait **quelle** magnifique boule de cristal, Rosa

superbe

partie **bizarre**

était **devenue** à la suite du vol, la **poule** de toutes les voyantes de la ville.

triste **risée**

Notre **alors**

Sa boule de cristal ne lui avait **pas** permis de prédire le cambriolage de

Pourtant **été**

appartement

son **arbre** .

voler

chatouillée **il**
Et cet homme, qui l'avait, **boule** était là devant elle : petit, fluet **ou** chauve.
ridiculisée **et**

blanche **attitude**
Il lui tendait sa boule, **avec** ne cessant de s'excuser pour «l'**emprunt** de sa
en **arrivé**

donc
boule». Il lui expliqua **s'** il était chômeur et qu'il avait **volé** la boule de
qu'

les **découvrir**
cristal pour connaître **ses** prochains numéros du loto afin de **gagner** le pactole.
où **jetons**

s' **homme**
La voyante lui demanda **qu'** il avait vu ces numéros. L' **inspecteur** déclara que
à **heureux**

numéro **toute**
oui, il s'était vu **fort** ,mais malheureux. Sa maison était devenue **un** coffre-fort
riche **pire**

verrou **perdu**
et il n'osait plus **fermer** de peur des voleurs. Il avait **aimé** tous ses amis.
sortir **peur**

crié **enfin**
« Alors, j'ai **décidé** d'abandonner et de vous rapporter **sa** boule » dit-il.
effrayé **votre**

loin **bien**
La voyante regarda **dans** sa boule de cristal, elle vit **donc** cet homme allait très
parmi **que**

chien
prochainement retrouver un **travail** et être heureux avec ses amis.
drôle