



**HAL**  
open science

## New tools for diagnosing and remedying dyslexia

Marie Lubineau

► **To cite this version:**

Marie Lubineau. New tools for diagnosing and remedying dyslexia. Human health and pathology. Université Paris Cité, 2023. English. NNT : 2023UNIP7192 . tel-04646682

**HAL Id: tel-04646682**

**<https://theses.hal.science/tel-04646682>**

Submitted on 12 Jul 2024

**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.

Université Paris Cité

École doctorale n°474 FIRE

*Laboratoire INSERM-CEA Cognitive Neuroimaging unit  
CEA/SAC/JOLIOT/Neurospin*

# New tools for diagnosing and remediating dyslexia

Par **Marie LUBINEAU**

Thèse de doctorat de sciences cognitives, psychologie,  
linguistique, philosophie de la pensée

Dirigée par **Stanislas DEHAENE**

Présentée et soutenue **publiquement**

le 23/11/2023

Devant un jury composé de :

**Stanislas DEHAENE**, Professeur HDR, Université Paris Saclay, Directeur de thèse

**Naama FRIEDMANN**, Professeur, Université de Tel Aviv, rapporteur

**Sylviane VALDOIS**, HDR, Université Grenoble Alpes, rapporteur

**Irène ALTARELLI**, Maître de conférence, Université Paris Cité, examinatrice

**Caroline HURON**, HDR, Université Paris Cité, examinatrice

**Johannes ZIEGLER**, Professeur HDR, Université d'Aix Marseille, examinateur

**Hervé GLASEL**, Président du CERENE, co-encadrant, membre invité

**Cassandra POTIER WATKINS**, Chargée de recherche, Collège de France, co-encadrante, membre invité

# TABLE OF CONTENT

<b>ACKNOWLEDGEMENTS</b> .....	5
<b>STATEMENT OF ORIGINALITY</b> .....	7
<b>SUMMARY</b> .....	8
<b>RESUME</b> .....	10
<b>RESUME SUBSTANTIEL</b> .....	12
<b>PREAMBLE</b> .....	17
<b>INTRODUCTION</b> .....	19
<b>READING AND ITS DEVELOPMENT</b> .....	19
Behavioral Reading Proficiency in Skilled Adult Readers .....	19
A model to explain behavioural results: the dual-route model of reading.....	23
Reading development in children .....	26
Brain networks for reading.....	29
How do these networks develop during reading acquisition?.....	31
How does language affect these observations?.....	32
<b>READING IN DYSLEXICS</b> .....	35
How do behavioural results and brain activity vary with dyslexia? .....	35
Early categorization of dyslexia profiles.....	38
Theories on the cognitive basis of dyslexia .....	40
Dyslexia remediation.....	47
<b>GOAL OF THE THESIS: NEW TOOLS FOR DIAGNOSING AND REMEDYING DYSLEXIA</b> .....	49
<b>CHAPTER 1 – EXAMINING THE IMPACT OF READING FLUENCY ON LEXICAL DECISION RESULTS ON FRENCH 6TH GRADERS</b> .....	51
<b>ABSTRACT</b> .....	52
<b>INTRODUCTION</b> .....	53
How do lexicality, length and frequency effects vary within a grade? .....	55
How is response to pseudowords affected by reading level?.....	55
<b>METHODS</b> .....	56
Participants.....	56
Reading fluency test .....	57
Lexical Decision Task .....	58
Data Analyses .....	61

<b>RESULTS</b> .....	62
Length and lexicality effects.....	64
Frequency effect on words.....	65
Interaction between length and frequency .....	67
Processing of pseudowords.....	69
<b>DISCUSSION</b> .....	71
<b>CONCLUSION</b> .....	76
<b>CHAPTER 2 – THE MARIETTE, A SCREENING TEST FOR READING ERRORS IN PRIMARY SCHOOLS ....</b>	<b>77</b>
<b>ABSTRACT</b> .....	78
<b>INTRODUCTION</b> .....	79
<b>METHOD</b> .....	85
Design of the Mariette .....	85
Participants.....	86
Procedure .....	87
Error characterisation.....	88
Other observables .....	91
<b>RESULTS</b> .....	92
Reading speed and success rate.....	92
Success rates for different word categories.....	93
Error rate for each error type.....	95
Other observables .....	97
Validity of the Mariette .....	99
<b>DISCUSSION</b> .....	100
<b>CONCLUSION</b> .....	103
<b>CHAPTER 3 – DOES WORD FLICKERING IMPROVE READING? NEGATIVE EVIDENCE FROM FOUR EXPERIMENTS USING LOW AND HIGH FREQUENCIES .....</b>	<b>105</b>
<b>ABSTRACT</b> .....	106
<b>INTRODUCTION</b> .....	107
<b>EXPERIMENT 1. IMPACT OF LOW-FREQUENCY FLICKERING ON NORMAL ADULT READERS .....</b>	<b>111</b>
Method.....	111
Results .....	114
<b>EXPERIMENT 2 – IMPACT OF LOW FREQUENCY FLICKERING ON DYSLEXIC CHILDREN .....</b>	<b>117</b>
Method.....	117
Results .....	118

<b>EXPERIMENT 3 – IMPACT OF HIGH FREQUENCY FLICKERING ON A GROUP OF DYSLEXIC STUDENTS</b>	120
Method .....	120
Results .....	124
<b>EXPERIMENT 4 - SINGLE CASE STUDY</b> .....	126
Method .....	126
Results .....	130
<b>GENERAL DISCUSSION</b> .....	133
<b>CONCLUSION AND FUTURE RESEARCH DIRECTIONS</b> .....	136
<b>GENERAL CONCLUSION</b> .....	137
Contributions to the science of reading .....	137
Contributions to the science of dyslexia .....	139
Our deliverables .....	140
<b>FUTURE RESEARCH DIRECTIONS</b> .....	140
Continuing to develop a battery of tests focusing on children's specific errors .....	140
Large scale studies to measure the prevalence of each of these types of dyslexia in French and in other languages .....	141
Find personalised solutions based on the dyslexia profile .....	141
<b>REFERENCES</b> .....	143
<b>APPENDIX</b> .....	161
Appendix 1. La Mariette .....	161

## ACKNOWLEDGEMENTS

Because a piece of research is never the work of an isolated individual, but rather the fruit of collaboration, I would like to take this opportunity to thank all the people who have supported me over the last three years. First and foremost, my thesis supervisor, Professor Stanislas Dehaene, who has guided me through these various projects. Working alongside him taught me rigour. He passed on to me his passion for cognitive science and education and enabled me to carry out meaningful work. I would also like to thank Hervé Glasel, Director of the CERENE schools, for funding this work and for his insight into my research. He gave me a better understanding of how the clinical world works, so that this work could then be applied. Special thanks to Cassandra Potier-Watkins, my mentor since I started at NeuroSpin, with whom I share this desire to improve the education system. Always available to answer my questions, she introduced me to experimental research and helped me to improve my English enormously!

This work was reviewed, commented on and criticised by my supportive thesis committee, including Lilliane Sprenger-Charolles and Michèle Mazeau. Their expert eye helped make this work better and steer it in the right direction. I would also like to thank all the schools that welcomed me for my various projects and enabled me to meet nearly a thousand students over the course of these three years. Thank you to my interns, Clémence, Audrey, Emma and Margaux, for their invaluable help in acquiring data for the experiments.

Thank you to all my colleagues in the laboratory, and in particular to Marie, Séverine, François, Lucas and the whole baby lab team for our fruitful meetings, Mathias for his help in understanding statistics and introducing me to web development, Vanna for her help with all the administrative formalities and her smile, Lorenzo and Caroline for our rich discussions about work and life, their invaluable advice and their friendship. Thanks also to Elena, Amélie, Camille and Nathalie, my colleagues at CERENE, who

opened me the doors of the diagnostic centre and helped me to better understand the link between research and clinic.

Last but not least, I wanted to thank my family and friends, who have accompanied me along the way, supporting and encouraging me through the good times and the not-so-good. A special thank to Antony, my team-mate with whom we've made progress on our life projects, and to my parents, my sister and my brother for the wonderful times we spend together, which make my everyday life so much brighter.

# STATEMENT OF ORIGINALITY

This PhD work has been conducted by myself, under the supervision of Stanislas Dehaene, and in collaboration with Cassandra Potier Watkins and Hervé Glasel. Each chapter of this work has been accepted or is currently in preparation for publication:

## Chapter 1

Lubineau, M., Potier Watkins, C. P., Glasel, H., & Dehaene, S. (2023). *Why do young readers vary in reading fluency? The impact of word length and frequency in French 6th graders* [Preprint]. Neuroscience. <https://doi.org/10.1101/2023.01.30.526188>, under review for OpenMind journal

## Chapter 2

Lubineau, M., Dehaene, S., Glasel, H., Potier Watkins, C. La Mariette, un test de dépistage des erreurs au primaire. In preparation

## Chapter 3

Lubineau, M., Potier Watkins, C. P., Glasel, H., & Dehaene, S. (2023). *Does word flickering improve reading? Negative evidence from four experiments using low and high frequencies*, Proceedings of the Royal Society B, <https://doi.org/10.1098/rspb.2023.1665> accepted in *Proceedings of the Royal Society B: Biological Sciences*



## SUMMARY

This thesis is a CIFRE project carried out in collaboration with CERENE, with the aim of developing and evaluating tools for diagnosing and remedying dyslexia, based on the solid literature on the dual route model of reading and the existence of different subtypes of dyslexia.

The CERENE welcomes a majority of pupils at the start of 6th grade. To provide schools with an effective tool for characterising the reading difficulties of their pupils, we first present a lexical decision test. Normed on 1501 pupils in 6<sup>th</sup> grade, this test also made it possible to study from a theoretical point of view the impact of reading fluency level on lexical decision results among pupils at the same school level. The 120 stimuli seen by each pupil varied in terms of lexicality, length (from 4 to 8 letters) and frequency (4 frequency bands, from very frequent to very rare). The pseudowords included spelling traps, transpositions and mirror substitutions between the letters bd and pq. We found that the effect of word length decreased with reading fluency, with the most struggling students showing a strong length effect even for frequent words. The effect of frequency also varied with fluency, but only in proportion to overall reading speed, suggesting that frequency affects the decision phase in all readers, while length has a greater impact on poor readers. Performance is also influenced by the type of pseudoword, and these effects vary little with fluency level.

Since it is also important to be able to detect specific reading deficits in young children, we then designed the Mariette, a test that takes into account the exact errors made by the child, unlike others such as the Alouette, which is widely used by researchers and practitioners to quantify reading fluency. It is a non-meaningful text containing regular words, irregular words, pseudowords and words that may lead to transposition or inter-word migration errors. The errors of 797 pupils from 1<sup>st</sup> to 5<sup>th</sup> grade were characterised in order to establish error thresholds above which a child is likely to present a specific deficit.

Finally, CERENE notes that more and more parents are asking about flickering reading aids for dyslexic children, such as the Lexilight® lamp or Lexilens® glasses. Since a few isolated cases of dyslexics helped by low-frequency flickering have been described in the literature, we first conducted two studies involving a computerised lexical decision task with a constant display or low-frequency flickering (10 or 15 Hz). In 375 normal-reading adults, flickering slowed down word recognition and slightly biased the decision towards pseudowords, while no effect was observed in 20 dyslexic students. Given the lack of data on flickering devices, we then assessed the impact of the Lexilight® lamp and Lexilens® glasses, which operate at higher frequencies, on reading fluency and letter identification in 22 dyslexic students. No detectable impact was observed. Finally, in the case of two participants who claimed to benefit from Lexilens® glasses, we orthogonally manipulated the fact that the glasses were actually switched on and the fact that the patient believed they were. Only a small placebo effect was observed in one of the participants. These results contrast sharply with marketing claims that these tools can help 90% of dyslexics.

As all these results were established with a view to practical applications, the major conclusions will also be the subject of short publications in French.

Keywords : reading, dyslexia, children, remediation

## RESUME

Cette thèse est un projet CIFRE mené en collaboration avec le CERENE, afin de développer et d'évaluer des outils de diagnostic et de remédiation de la dyslexie, élaborés sur la base de la solide littérature sur le modèle de double voie de la lecture et l'existence de différents sous-types de dyslexie.

Le CERENE accueille une majorité d'élèves à l'entrée en 6<sup>ème</sup>. Pour fournir aux écoles un outil efficace pour caractériser les difficultés de lecture de leurs élèves, nous présentons dans un premier temps un test de décision lexicale. Normé sur 1501 élèves de 6ème, ce test a également permis d'étudier du point de vue théorique l'impact du niveau de fluence de lecture sur les résultats de la décision lexicale chez des élèves d'un même niveau scolaire. Les 120 stimuli vus par chaque élève variaient en termes de lexicalité, de longueur (de 4 à 8 lettres) et de fréquence (4 bandes de fréquence, de très fréquent à très rare). Les pseudo-mots comprenaient des pièges orthographiques, des transpositions et des substitutions miroir entre les lettres bd et pq. Nous avons constaté que l'effet de la longueur des mots diminuait avec la fluence de lecture, les élèves les plus en difficulté montrant un fort effet de longueur même pour les mots fréquents. L'effet de la fréquence variait aussi avec la fluence, mais uniquement en proportion de la vitesse de lecture globale, ce qui suggère que la fréquence affecte la phase de décision chez tous les lecteurs, tandis que la longueur impacte davantage les mauvais lecteurs. Les performances sont également influencées par le type de pseudo-mot, et ces effets varient peu avec le niveau de fluence.

Puisqu'il est aussi important de pouvoir détecter des déficits de lecture spécifique chez les jeunes enfants, nous avons ensuite conçu la Mariette, un test qui tient compte des erreurs exactes faites par l'enfant contrairement à d'autres comme l'Alouette, pourtant largement utilisé par chercheurs et praticiens pour quantifier la fluidité de la lecture. C'est un texte non signifiant contenant des mots réguliers, des mots irréguliers, des pseudo-mots et des mots pouvant entraîner des erreurs de transposition ou de migration inter-mots. Les erreurs de 797 élèves du CP au CM2 ont été caractérisées

afin d'établir des seuils d'erreurs au-delà desquels un enfant est susceptible de présenter un déficit spécifique.

Enfin, le CERENE constate que de plus en plus de parents s'interrogent sur les outils d'aide à la lecture pour enfants dyslexiques fondés sur la stroboscopie comme la lampe Lexilight® ou les lunettes Lexilens®. Puisque quelques cas isolés de dyslexiques aidés par le clignotement à basse fréquence ont été décrits dans la littérature, nous avons d'abord mené deux études impliquant une tâche de décision lexicale informatisée avec un affichage constant ou un clignotement à basse fréquence (10 ou 15 Hz). Chez 375 adultes normaux-lecteurs, le clignotement a ralenti la reconnaissance des mots et biaisé légèrement la décision en faveur des pseudo-mots, tandis qu'aucun effet n'a été observé chez 20 élèves dyslexiques. Compte tenu du manque de données sur les dispositifs stroboscopiques, nous avons ensuite évalué l'impact de la lampe Lexilight® et des lunettes Lexilens®, qui fonctionnent à des fréquences plus élevées, sur la fluidité de la lecture et l'identification de lettres chez 22 élèves dyslexiques. Aucun impact détectable n'a été observé. Enfin, chez deux participantes qui disaient bénéficier des lunettes Lexilens®, nous avons orthogonalement manipulé le fait que les lunettes soient effectivement allumées et le fait que la patiente croit qu'elles le soient. Seul un petit effet placebo a été observé chez l'une des participantes. Ces résultats contrastent fortement avec les affirmations marketing selon lesquelles ces outils peuvent aider 90 % des dyslexiques.

Tous ces résultats ayant été établis en vue d'applications pratiques, les conclusions majeures feront également l'objet de courtes publications en français.

Mots clés : lecture, dyslexie, enfants, remédiation

## RESUME SUBSTANTIEL

Cette thèse est un projet CIFRE, mené en collaboration avec les écoles et le centre d'évaluation neuropsychologique du CERENE, afin de développer et d'évaluer des outils de diagnostic et de remédiation de la dyslexie. Ces outils ont été développés sur la base de la solide littérature existante sur le modèle double voie de la lecture (Coltheart et al., 2001) selon lequel la lecture résulte d'une compétition entre deux voies de traitement des informations écrites. D'un côté, la voie lexicale ou voie d'adressage, reconnaît les mots déjà stockés dans le lexique. Ce lexique est organisé par fréquence, rendant les mots plus fréquents plus rapides à extraire que les mots moins fréquents. Cette voie permet une lecture rapide des mots déjà connus du lecteur. D'un autre côté, la voie sublexicale ou voie d'assemblage, permet le déchiffrement de mots nouveaux ou de pseudo-mots via l'application des règles de conversion grapho-phonémiques. Cette voie est plus lente, et son utilisation est marquée par un fort effet de longueur (Coltheart, 2007). L'entrée dans ces deux voies est précédée d'une analyse visuo-orthographique, au cours de laquelle le lecteur va identifier les lettres composant le mot, et les associer à leur position au sein du mot ainsi qu'entre les mots voisins, dans le cas de lecture de textes par exemple (Friedmann & Coltheart, 2018). Une prédiction du modèle doubles voies est qu'en plus des dyslexies phonologique, résultant d'un déficit de la voie sublexicale, et de surface, résultant d'un déficit de la voie lexicale, un certain nombre de dyslexies pourraient émerger de déficits au niveau de l'analyse visuo-orthographique préalable à ces deux voies (Friedmann & Coltheart, 2018). De fait, au moins deux grandes catégories de dyslexies ont été identifiées à ce niveau : la dyslexie de position de lettres, et la dyslexie attentionnelle. La dyslexie de position de lettres affecte sélectivement la capacité d'encoder la position relative des lettres dans les mots. Les migrations de lettres apparaissent le plus souvent entre deux lettres intérieures et adjacentes, particulièrement lorsque l'erreur forme elle aussi un mot – ainsi, « fiable » est lu *faible*, ou « magner » *manger*. Ces erreurs se produisent à la fois en lecture à haute voix et silencieuse, dans n'importe quelle tâche qui implique d'identifier ou de comprendre un mot. Des cas de pure dyslexie de position de lettres, ne présentant aucun déficit

phonologique, ont été identifiés dans un grand nombre de langues (Friedmann & Haddad-Hanna, 2012; Friedmann & Rahamim, 2007; Güven & Friedmann, 2019; Kohnen et al., 2012; Potier Watkins et al., 2023). La dyslexie attentionnelle est une autre dyslexie visuelle qui altère sélectivement la capacité d'attacher une lettre au mot d'où elle provient. Dans ce type de dyslexie, les lettres sont correctement identifiées et conservent presque toujours leur position relative, mais elles migrent vers un mot voisin – par exemple « suis pois » est lu *puis pois*. Les migrations peuvent avoir lieu à la fois horizontalement et verticalement, apparaissent le plus souvent en fin de mot, et en français, plus fréquemment du deuxième mot vers le premier (ce qui indique qu'il ne s'agit pas simplement de persévérations d'une lettre ou d'un phonème). Elles se manifestent à la fois par des substitutions, des additions ou des omissions de lettres présentes dans les mots voisins (Friedmann et al., 2010; Potier Watkins et al., 2023; Rayner et al., 1989). Les outils que nous avons développés au cours de cette thèse s'appuient sur cette théorie.

Le CERENE accueille une majorité d'élèves à l'entrée en 6ème. Dans notre **premier chapitre**, nous décrivons la tâche de décision lexicale que nous avons développée pour ce niveau, afin de fournir aux écoles un outil efficace pour caractériser les difficultés de lecture rencontrées par les élèves. Normée sur 1501 élèves de 6ème représentatifs de la population française, cette épreuve a également permis d'étudier, d'un point de vue plus théorique, l'impact du niveau de fluence de lecture sur les résultats de la décision lexicale chez des élèves de même niveau scolaire. Nous avons testé une prédiction du modèle à double voie : à mesure que la fluence augmente, les variations dans les résultats devraient refléter une diminution de la dépendance au décodage et une augmentation de l'utilisation de la voie lexicale. Les 120 stimuli vus par chaque élève, tirés au hasard d'une base de données de plus de 3000 mots et de plus de 1000 pseudo-mots, variaient en termes de lexicalité (mots ou pseudo mots), de longueur (de 4 à 8 lettres) et de fréquence (4 bandes de fréquence, de très fréquent à très rare). Les pseudo-mots comprenaient notamment des pièges orthographiques (« bage », « plase »), des transpositions (« pafrois », « regadrer ») et des substitutions en miroir entre les lettres bd et pq (« aibe », « musipue »), ainsi que des contrôles pour chaque catégorie. Comme prévu, nous avons

constaté que l'effet de la longueur des mots variait considérablement en fonction de la fluence de la lecture, les élèves les plus en difficulté montrant un fort effet de longueur même pour les mots fréquents. L'effet de la fréquence varie également avec la fluence, mais uniquement en proportion de la vitesse de lecture globale, ce qui suggère que la fréquence affecte la phase de décision chez tous les lecteurs, tandis que la longueur a un impact plus important sur les mauvais lecteurs. Les temps de réponse et les erreurs ont également été influencés par le type de pseudo-mot, mais ces effets varient très peu avec le niveau de fluence.

S'il est important de pouvoir évaluer la lecture avec précision et rapidité à l'entrée en 6ème, il est également essentiel de pouvoir détecter des déficits spécifiques chez les enfants plus jeunes. Le test de l'Alouette (Lefavrais, 2005), bien que largement utilisé par les praticiens et la communauté scientifique pour quantifier la fluidité de la lecture, ne fournit que des mesures quantitatives de la vitesse et de la précision globales de la lecture, et ne tient pas compte des erreurs exactes commises par l'enfant. Or pour détecter des troubles spécifiques tels que des dyslexies de position de lettres ou des dyslexies attentionnelles, il est essentiel de s'intéresser aux erreurs exactes de l'enfant et d'inclure dans les tests des stimuli piégeux pour faire en sorte que chaque types d'erreur puisse se manifester – chez certains patients par exemple, seule la présentation de mots « transposables », comme fiable ou magner, met en évidence des erreurs de position de lettres qui, sans cela, passeraient inaperçues, la dyslexie se manifestant alors seulement par un ralentissement de la lecture, voir une absence de déficit aux tests standards (Kohnen et al., 2012; Potier Watkins et al., 2023). Afin de doter le CERENE d'un test plus précis et tout aussi efficace, le **deuxième chapitre** de ce travail présente la Mariette, un texte non signifiant de 294 mots contenant notamment des mots réguliers, des mots irréguliers, des pseudo-mots, des mots pouvant conduire à des transpositions (cirque/crique) ainsi que des paires de mots pouvant conduire à des migrations inter-mots (cape page / cape cage). Les erreurs de 797 élèves du CP au CM2 ont été caractérisées et des seuils d'erreurs au-delà desquels un enfant est susceptible de présenter un déficit spécifique tel qu'une dyslexie de position des lettres ou une dyslexie attentionnelle ont été établis pour chaque niveau. Un test comme celui-ci permet également d'évaluer

la lecture de l'enfant dans les conditions les plus écologiques possibles car contrairement à une liste de mots, la lecture d'un texte nécessite une bonne utilisation des mécanismes oculomoteurs.

Enfin, le CERENE constate que de plus en plus de parents s'interrogent voire achètent des outils d'aide à la lecture pour enfants dyslexiques fondés sur la stroboscopie, comme la lampe Lexilight® ou les lunettes Lexilens®. Le fondement scientifique derrière cette idée semble pourtant extrêmement mince. Il découle d'une étude de Le Floch et Ropars (2017), qui affirme que la dyslexie est causée par une anomalie rétinienne entraînant la formation d'images miroirs illusoires et se traduisant par l'absence d'un œil dominant, à laquelle il serait possible de remédier par un clignotement à haute fréquence. La logique de cette étude est très discutable : aucun score de lecture n'a été fourni dans cette étude ; une anomalie rétinienne, si elle était correctement documentée, n'expliquerait pas les dissociations observées dans la dyslexie, par exemple entre la lecture des chiffres et des lettres (Dotan & Friedmann, 2019 ; Friedmann, Dotan, et al., 2010) ; enfin, à notre connaissance, les résultats de cette étude n'ont jamais été répliqués. Il serait donc facile de faire de ce clignotement une proposition excentrique, s'il n'y avait pas plusieurs contre-arguments possibles. Tout d'abord, les fabricants semblent réussir à vendre leurs produits et finalement, quelques études ont décrit des adultes souffrant de troubles de la lecture en miroir qui ont été aidés par un clignotement à basse fréquence (McCloskey et al., 1995 ; McCloskey & Rapp, 2000a ; Pflugshaupt et al., 2007 ; Vannuscorps et al., 2022). Dans le **troisième chapitre** de ce travail, nous avons donc cherché à mieux comprendre si le clignotement à basse ou haute fréquence pouvait faciliter la lecture pour les lecteurs normaux et les dyslexiques. Nous avons tout d'abord mené deux expériences impliquant une tâche de décision lexicale informatisée avec un affichage constant ou un clignotement à basse fréquence (10 ou 15 Hz). Chez 375 adultes normaux-lecteurs, le clignotement a sensiblement ralenti la reconnaissance des mots, tout en biaisant légèrement la décision en faveur des pseudo-mots, tandis qu'aucun effet significatif n'a été observé chez 20 enfants dyslexiques. Nous nous sommes ensuite intéressés à l'impact du clignotement à haute fréquence. Pour ce faire, nous avons utilisé un environnement plus naturel (lecture sur papier) ainsi que la lampe et les lunettes décrites ci-dessus. Nous avons testé une



fréquence de clignotement de ~80Hz, sur un groupe d'enfants dyslexiques qui ne connaissaient pas ces dispositifs, puis sur deux patientes, une adulte et une enfant, qui affirmaient toutes deux être aidées par les lunettes au quotidien. Aucun impact détectable n'a été observé sur le groupe de 22 enfants dyslexiques. Enfin, chez les deux participantes qui prétendaient bénéficier de lunettes, nous avons orthogonalement manipulé le fait que les lunettes soient effectivement allumées et le fait que la patiente croit qu'elles le soient. Seul un petit effet placebo a été observé chez l'une des participantes. Nos résultats contrastent fortement avec les affirmations marketing selon lesquelles ces outils peuvent aider 90 % des dyslexiques, et soulignent le rôle d'une recherche scientifique rigoureuse pour permettre aux personnes dyslexiques de prendre des décisions en connaissance de cause.

L'ensemble de ces travaux a été réalisé dans le but de fournir au CERENE et, plus largement, aux praticiens et aux enseignants, des solutions applicables pour le dépistage et la caractérisation des difficultés de lecture. Deux outils ont été conçus dans ce but. La décision lexicale a été standardisée pour les élèves de 6<sup>ème</sup> et est désormais accessible gratuitement en ligne, sur ordinateur et téléphone portable (<https://neurospin-data.cea.fr/exp/PiegeAmots/>). A la fin de chaque test, d'une durée d'environ 3 minutes, un feedback automatique est généré, indiquant les points forts et les points faibles de l'élève. L'utilisation régulière du test permet de voir l'évolution de l'élève et de mesurer ses progrès. Le Mariette a été standardisée pour les élèves du CP au CM2 et sera bientôt mise à la disposition des enseignants et des praticiens qui souhaitent l'utiliser. Un guide d'utilisation comprenant des normes et des règles pour faciliter le repérage des erreurs accompagnera le test. Enfin, avec notre étude sur les outils d'aide à la lecture fondés sur le principe de la stroboscopie, nous avons fourni aux professionnels paramédicaux du CERENE des arguments pour répondre aux questions des parents sur les outils de lecture qui sont actuellement sur le marché avec la promesse infondée d'aider près de 90% des enfants dyslexiques.

## PREAMBLE

Developmental dyslexia is a learning disability that affects between 3% and 12% of the population, depending on language and definition (Di Folco et al., 2022; Lindgren et al., 1985; Roongpraiwan et al., 2002; Yang et al., 2022). It is characterized by a range of reading difficulties, which can vary from extreme slowness in reading to a complete inability to recognize letters. Dyslexia is part of a broader spectrum of neurodevelopmental disorders, including dyspraxia, dysphasia, and dyscalculia, all of which significantly impact children's ability to learn in school.

While initial signals to difficulty in reading may raise concerns as early as the end of kindergarten, formal diagnosis usually occurs at the end of the 2nd grade (Mazeau et al., 2022). For this reason, dyslexia is typically defined by a delay of approximately two years in the development of reading skills. Diagnosis is typically done through assessments by either neuropsychologists or speech therapists and followed up with a tailored program of personalized support depending on the severity of the disorder. However, dyslexia is not a condition that is 'cured'. Therapy is usually in the goal, for milder cases, to provide the child with the additional focused practice to help them learn to read or help the child learn to use tools like audio reading aids and spell checkers to counter their difficulties with reading. In the latter case of more mild dyslexia, the deficit does not prevent reading but persists as slow reading, while in the later, more severe case, reading remains difficult.

In recent years, specialized schools dedicated to providing optimal support for these children have emerged throughout France. Among the pioneers is the school CERENE (<http://cerene-education.fr>), which currently serves 400 students, with 70% of them being dyslexic. The goal of the school is to accommodate students from 3rd to 9th grade with normal intelligence, but with a 'dys' profile that makes attending normal school difficult. CERENE provides personalized teaching through small-class settings, equipping students with laptop computer loaded with software tools such as spellcheckers and voice feedback, training with occupational therapists to learn how to use IT tools, and

psychologists, neuropsychologists, and speech therapists to provide on-site support. The school's primary objective is to prepare students for a successful return to a standard curriculum with appropriate adaptations once they have completed their time at CERENE. In other words, CERENE allows students to keep up with their school learning, while also learning how to integrate tools that allow them to navigate life-long learning despite their 'dys' related difficulties.

In addition to the school, CERENE also has a neuropsychological diagnostic centre where neuropsychological assessments are carried out on children from kindergarten to secondary school. These assessments evaluate a wide variety of cognitive dimensions, including written language, enabling a diagnosis to be made and recommendations for adapting the child's schooling and managing the disorder. Depending on the profile, these children may be referred to a CERENE school.

This PhD is the result of a collaboration with CERENE, the aim of which was to improve dyslexia screening in France and to evaluate the effectiveness of remediation tools already on the market. The tools developed and the evaluation protocols are based on the vast literature available in the field of reading and dyslexia, an overview of which is given in this introduction.

# INTRODUCTION

Reading, a relatively recent human faculty that emerged alongside writing approximately four millennia before our era, represents a pivotal moment in the shift from prehistory to history. It serves as the bridge between oral language and the written code. The ability to read signifies the acquisition of a mental information processing system capable of navigating the transition from text to spoken language. This process is intricate, demanding a cognitively intensive learning phase and resulting in alterations to the organization of the brain. In this discussion, we will explore the various aspects of reading, encompassing proficient adult readers, the challenges of dyslexia, and reading development in the child's brain.

## READING AND ITS DEVELOPMENT

### **Behavioral Reading Proficiency in Skilled Adult Readers**

In the realm of skilled adult readers, words seem to flow towards understanding. We effortlessly identify signs during a leisurely stroll, news headlines immediately pique our interest as we scroll, and the words in books come alive in our thoughts. Words are everywhere, and grasping their meaning feels as natural as breathing. It is only when an expert reader encounters an unfamiliar word, such as 'Sesquipedalian' (meaning the use of long words or verbosity), that a brief interruption occurs in their thought process. In these moments, they may adopt a more deliberate approach, akin to a child, by slowly sounding out and listening to the word to fully comprehend it. Two reading tasks have emerged as particularly influential in shedding light on the cognitive processes underlying both effortless and effortful reading. They reveal that even in the case of seemingly effortless reading, there are intricate cognitive processes at work that can be understood and manipulated. These tasks are the "Naming" and "Lexical Decision" tasks. This subset is deliberately "reductive" in order to limit it to isolated word reading tasks, essentially reading without the influence of contextual cues.

In the naming task, participants are presented with visually displayed words (e.g., "black") or pseudowords (e.g., "bluck") and are asked to name them promptly and accurately. On the other hand, the lexical decision task assesses silent reading processes. Participants are shown a series of letter strings, which can be real words or pseudowords, and must quickly determine whether each item is a real word or not, often by pressing a designated key as rapidly as possible. In these tasks, researchers typically analyze participants' response times and error rates. Both tasks are commonly employed in psychological experiments to investigate the speed and accuracy of lexical access, shedding light on how the brain processes and recognizes words. These paradigms reveal various effects, with the most significant ones pertaining to the lexicality of stimulus (real word or pseudoword), word frequency, and word length, which we will introduce below, and which are illustrated in Figure 1.

Several studies describe the presence of a lexicality effect: reading times for non-words are longer than those for regular words (B. S. Weekes, 1997; Seidenberg et al., 1984; Forster & Chambers, 1973; Juphard et al., 2004a). This lexicality effect interacts most of the time with that of length in the following way: for words between 4 and 8 letters, length does not influence reading performance once the neighbourhood effects<sup>1</sup> have been corrected, unlike the situation observed for pseudowords, for which reading time increases with length (Bijeljac-Babic et al., 2004a; B. S. Weekes, 1997), known as the length effect. Large-scale studies of several tens of thousands of words, conducted in French and English, have shown that the effect of word length follows a U-shape curve. It decreases between 2 and 4-5 letters, reaches a plateau between 4-5 and 8-9 letters and increases again for words of more than 9 letters. (Ferrand et al., 2010; New et al., 2006a).

Other studies have highlighted an effect of word frequency, with low-frequency words, which are rarer, taking longer to read than high-frequency words, which are much more frequently used

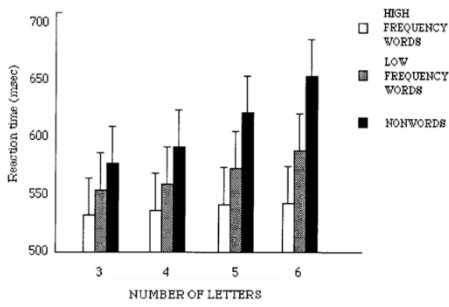
---

<sup>1</sup> The "neighborhood effect" refers to the influence or impact of words that are similar in spelling to the target word being evaluated.

(Andrews, 1992, 1989; Forster & Chambers, 1973; McCann & Besner, 1987; Seidenberg et al., 1984; Burani et al., 2002). This occurs despite understanding the non-frequent word. More precisely, the response times to the lexical decision vary linearly with the logarithm of the word frequency (Norris, 2006). Data from the English lexicon project also show that this parameter is the best predictor of lexical decision response times, and that its influence on naming results is more tenuous (Baayen et al., 2006; Balota et al., 2004; Cortese & Khanna, 2007). This frequency effect interacts with the effect of word regularity with irregular words read less quickly than regular words, particularly low-frequency words suffering the greatest increase in response time (Paap & Noel, 1991; Seidenberg et al., 1984).

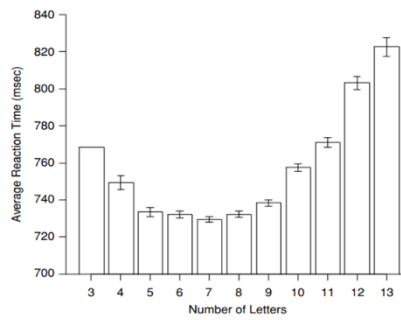
The number of neighbours is another factor that affects the recognition of written words. The neighbourhood effect refers to the influence or impact of words that are similar in spelling to the target word being evaluated. This is classically measured using Coltheart's N (Coltheart et al., 1977), which considers as "neighbours" two strings of the same length that differ by only one letter. The effect of neighbourhood interacts with that of frequency in the following way: word reading times are not affected by the number N of neighbours when words are of high frequency, but there is a facilitating effect of a large number of neighbours on lexical decision and naming when words are of low frequency (Andrews, 1989, 1992). This neighbourhood effect is reversed for the reading of pseudowords: the response to pseudowords is longer for high N pseudowords (e.g., the pseudoword 'plut' is neighbour to the frequent word 'plus') than for low N pseudowords (e.g., the pseudoword 'spowl' is neighbour to the infrequent word 'scowl') (Andrews, 1989; Coltheart et al., 1977).

**a. Interaction between lexicality, length and frequency in a naming task**

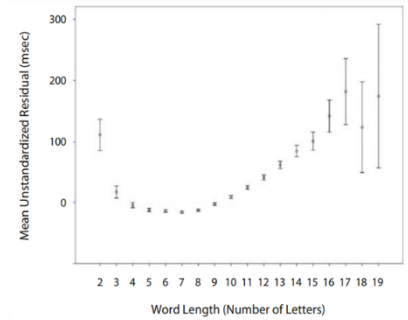


(Weekes et al. 1997)

**b. Length effect on a lexical decision task in English and in French**

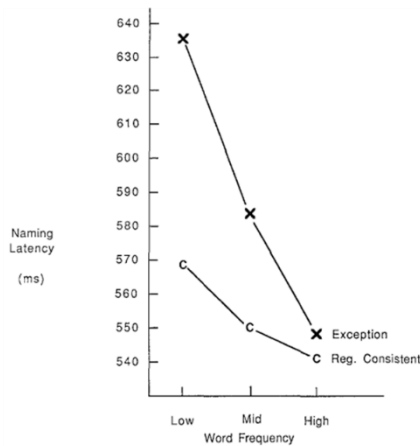


(New et al. 2006)



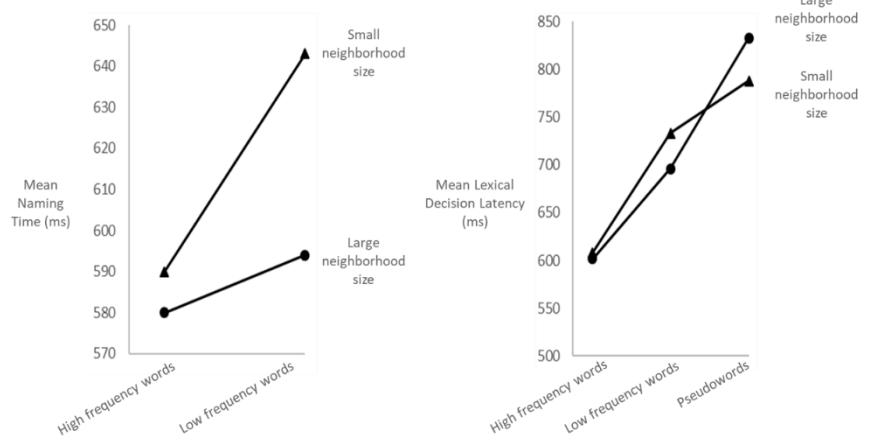
(Ferrand et al. 2010)

**c. Interaction between frequency and regularity in a naming task**



(Paap & Noel, 1991)

**d. Interaction between frequency and neighbourhood size in both lexical decision and naming**



(Andrews, 1989)

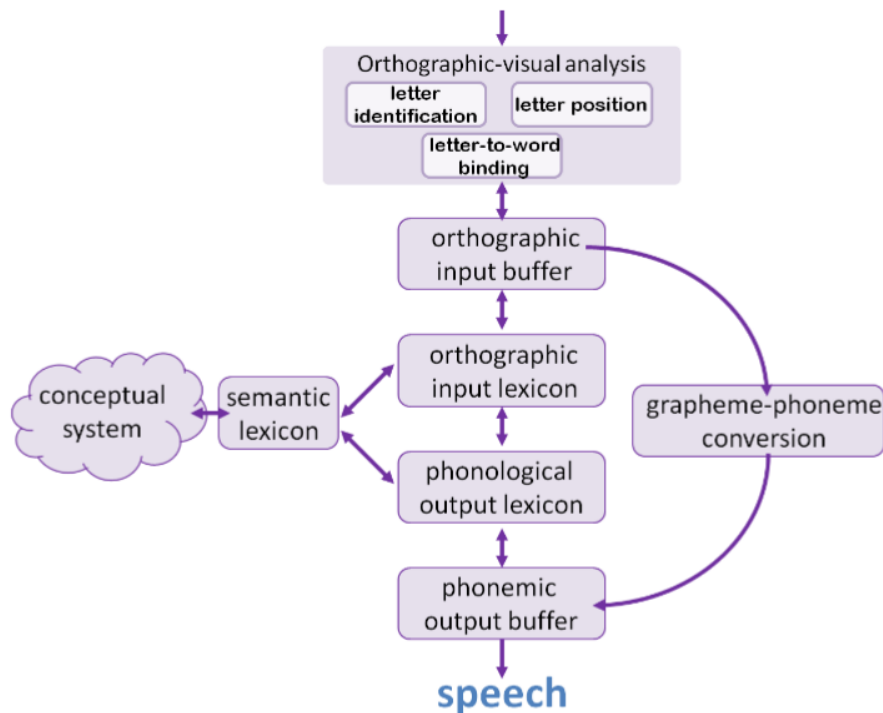
**Figure 1.** Illustration of the effects of lexicality, length, frequency, regularity and neighbourhood size and their interactions in adults. **(a) Interaction between lexicality, length and frequency in a naming task:** reading times are constant for frequent words no matter the length whereas they increase for less frequent words and even more for nonwords (Weekes et al. 1997). **(b) Length effect on a lexical decision task in English and in French:** it follows a U-shape curve, decreasing between 2 and 4-5 letters, reaching a plateau between 4-5 and 8-9 letters and increasing again for words of more than 9 letters (Ferrand et al., 2010; New et al., 2006a). **(c) Interaction between frequency and regularity in a naming task:** irregular words are read less quickly than regular words, particularly low-frequency words (Paap & Noel, 1991). **(d) Interaction between frequency and neighbourhood size in both lexical decision and naming:** the number N of neighbours does not affect word-reading times when words are of high frequency, but there is a facilitating effect of a large number of neighbours when words are of low frequency, an effect that is reversed for pseudowords (Andrews, 1989).

## **A model to explain behavioural results: the dual-route model of reading**

The results from research in the 1970s marked the emergence of the first models of reading. While various models have been extensively discussed in the literature, my work focused on Coltheart's Dual-Route model. The reason for this is quite clear: Coltheart's Dual-Route model is the most prevalent and influential model of reading. Its prominence arises from its substantial empirical backing, drawn from a diverse array of studies conducted in the fields of psycholinguistics, cognitive psychology, and, notably, research in reading development within education. Moreover, this model has shown remarkable predictive power in explaining a wide spectrum of reading behaviours observed in skilled readers when implemented (Coltheart et al., 2001a; Ziegler et al., 2014, 2020). Our inclination toward this model is also grounded in its lucidity and simplicity, which we will delve into further below.

This Dual-Route model describes the reading process as the outcome of two competing pathways. On the one hand, the lexical pathway accesses words already stored in the mental lexicon. This lexicon is organized by word frequency, such that 'more' frequent words are quicker to retrieve than less frequent ones. This pathway lends itself to fast fluent reading of known words. On the other hand, the sublexical pathway enables reading words that have not been stored in the mental lexicon, such as unknown words or pseudowords, via the application of grapheme-phoneme correspondence (GPC) rules (Coltheart, 2007). More recently, this model's explanation of central reading processes have been expanded on to include peripheral elements of reading, such as visual processes (Figure2). For a more complete view, see Friedmann and Coltheart in the Handbook of communication disorders (Friedmann & Coltheart, 2018).





**Figure 2.** The dual route model of reading (Friedmann & Coltheart, 2018)

Access to the centralized routes of lexical and sub-lexical reading is preceded by a visuo-orthographic analysis during which the reader identifies the letters that make up the word, orders them to avoid confusion between anagrams, and associates them with the word he or she is currently reading, and not with the following or preceding words. The information then enters the orthographic input buffer. It is probably at this level that morphological decomposition takes place, so that only the stems are sent to the lexical and sublexical pathways. Morphological affixes are processed in parallel and collected at the end of the pathway. The lexical pathway comprises several stages. The information first enters the orthographic input lexicon, which points to phonetic inputs, in the phonetic input lexicon, and semantic inputs, via the semantic lexicon, associated with the written form of the word when it is known to the reader. It is this semantic lexicon, linked to the conceptual system, which enables us to understand written words. At the same time, pronunciation is also activated. In the sublexical pathway, reading is done through grapheme-phoneme conversion. Conversion are

maintained in the phonological output buffer, where the word and its morphological affixes are recomposed allowing the reader to access pronunciation and meaning.

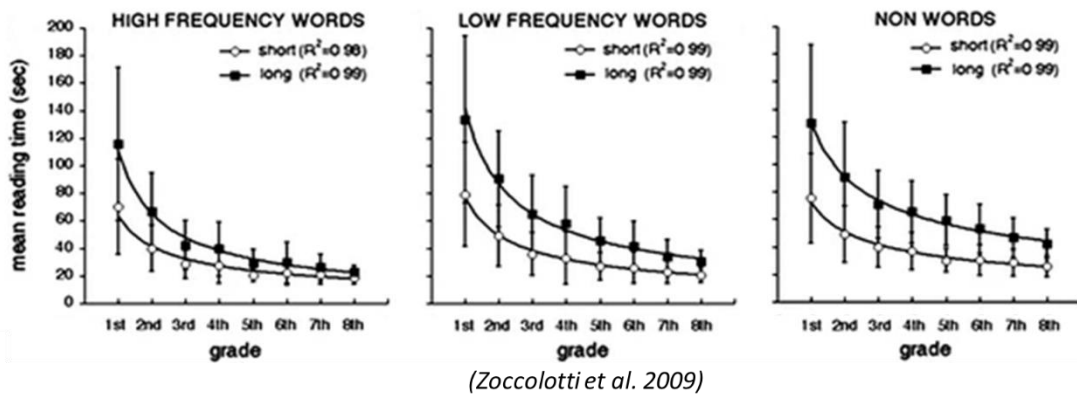
The dual-route model of reading provides a framework to understand many of the observed behavioral effects I have presented. For instance, the lexicality effect can be explained by the speed difference between the two reading pathways. When readers encounter known regular words, they can rapidly process them through the lexical route, which is the fastest route. However, when confronted with pseudowords, they must resort to the slower sublexical route. When reading via the lexical route, letters in medium-length words (around 4 to 8 letters) can be processed simultaneously. In contrast, the sublexical route requires sequential letter processing. Due to this difference, we do not observe a word length effect for medium-length words, but reading time for pseudowords increases linearly with the number of letters. The frequency effect is a consequence of how the mental lexicon is organized within the lexical route. In this mental lexicon, words are organized by their frequency of use, making high-frequency words more readily accessible than low-frequency ones. This organization influences how quickly we can recognize and process words during reading. Using the dual route model, we can also explain the regularity effect and its interaction with the frequency effect: frequent irregular words are rapidly decoded by the lexical route, just like regular words, unlike infrequent irregular words, for which the two routes generate contradictory information. Resolving this conflict requires additional processing time compared with reading infrequent regular words. In addition, infrequent irregular words will have a greater tendency to produce regularisations than frequent irregular words, as long as the previous conflict is resolved in favour of the sublexical route. Finally, the inversion of the effect of the number of neighbours, which is facilitative for words but deleterious for pseudowords, can be explained by interference from the lexical route, which interferes with the reading of pseudowords via the sublexical pathway by emitting possible neighbouring output words.

## Reading development in children

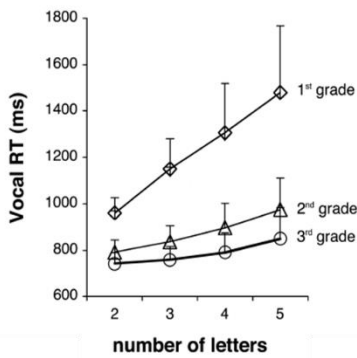
The effects of lexicality, length, frequency, and regularity evolve over the course of learning to read (Figure 3). It only takes a few months for the first lexicality effect to appear, which interacts with frequency: pupils in 1<sup>st</sup> grade read high-frequency words faster than pseudowords, but their reading time is the same between pseudowords and low-frequency words. This lexicality effect increases with age, appearing for low-frequency words from third grade onwards (Zoccolotti et al., 2009a). At the same time, the effect of word length, which is very present in young children, diminishes progressively with age and disappears once the fundamental mechanisms of reading have been acquired (Bijeljac-Babic et al., 2004a; Burani et al., 2002; Schröter & Schroeder, 2017; Zoccolotti et al., 2005a). This reduction is modulated by frequency, with the length effect fading more quickly for low-frequency words than for high-frequency words (Zoccolotti et al., 2009a). As in the case of adults, the length effect persists for pseudowords and its size remains constant during learning (Zoccolotti et al., 2009a).

The effect of frequency can also be observed from the first years of reading acquisition. It reaches his maximum effect in third grade, then decreases substantially, particularly for longer words. Unlike the length effect, it remains significant even for the most experienced readers (Schmalz et al., 2013a; Schröter & Schroeder, 2017; Zoccolotti et al., 2009a). Finally, the effect of frequency interacts with that of regularity in similar ways to adults, even in third grade: the effect of regularity is present for low-frequency words, but not for high-frequency words (Schmalz et al., 2013a).

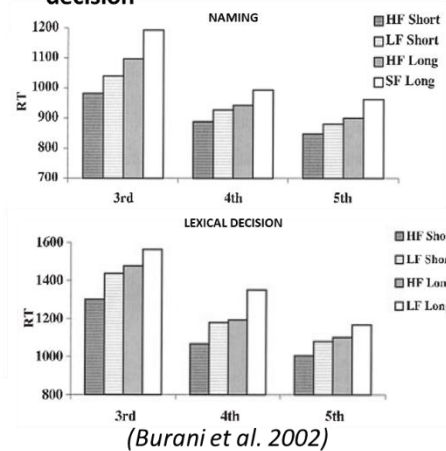
a. Interaction between lexicality, length and frequency in a naming task



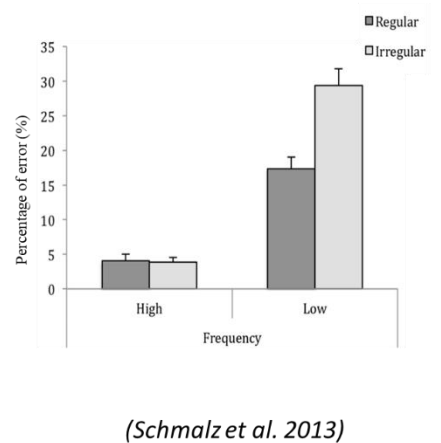
b. Diminution of the length effect between 1<sup>st</sup> and 3<sup>rd</sup> grade in a naming task



c. Interaction between length and frequency in both naming and lexical decision



d. Interaction between frequency and regularity in a lexical decision task



**Figure 3.** Illustration of the effects of lexicality, length, frequency, regularity and neighbourhood size and their interactions in the course of learning. **(a) Interaction between lexicality, length, and frequency in a naming task:** reading time difference between short and long words (length effect) decreases as a function of grade, faster for high frequency than low-frequency words. The length effect on pseudowords never disappears, even for the oldest students (Zoccolotti et al. 2009). **(b) Diminution of the length effect between 1<sup>st</sup> and 3<sup>rd</sup> grade in a naming task:** naming time linearly increases with length in 1<sup>st</sup> graders and remains constant in 3<sup>rd</sup> graders (Zoccolotti et al. 2005). **(c) Interaction between length and frequency in both naming and lexical decision:** the length effect on high frequency words that exists in 3<sup>rd</sup> graders almost fades in 5<sup>th</sup> graders, whereas is remains significant for low-frequency words (Burani et al. 2002). **(d) Interaction between frequency and regularity in a lexical decision task:** the effect of regularity is present for low-frequency words, but not for high-frequency words already in 3<sup>rd</sup> graders (Schmalz et al. 2013).

Through the lens of the dual-route model of reading, we can interpret these observations as follows. At the early stages of reading instruction, children begin by learning the rules of converting graphemes into corresponding phonemes. This fundamental step sets the foundation for the development of their sublexical pathway. During this initial phase, children predominantly rely on this sublexical route as

they gradually build their orthographic lexicon, which is essentially their mental repository of known words. This initial reliance on the sublexical pathway explains the substantial length effect often observed in young readers. In simpler terms, young readers tend to take more time to read longer words because they are primarily utilizing the sublexical route. As these readers gain experience and accumulate word knowledge, they become more proficient in recognizing words. This reduced dependence on the sublexical route leads to a decrease in the length effect. Furthermore, this decrease occurs more rapidly for high-frequency words compared to low-frequency words. This pattern reflects the stocking of words in the mental lexicon and the development of an experience-dependent lexical pathway. This interpretation aligns with the emergence of the frequency effect, which coincides with the decrease in the length effect. The frequency effect serves as an indicator of the utilization of the lexical pathway in reading.

Along with the central processing changes that happen in reading, learning to read also modifies peripheral aspects to fluency. For example, the visual system is modified with learning to read. The Dual route model explains this through several different visual processing stages. One of these processes, letter recognition, is particularly fastidious for letters that are symmetrical with each other (b, d, p and q, u and n, m and w), because of the mirror generalisation inherent in our visual system, which enables us in other contexts to associate the same name with an object and its symmetrical counterpart, for example. The ability to differentiate between these letters is very closely linked to the ability to unlearn mirror generalisation (Ahr et al., 2016). Mirror confusion is generally very common in young children and disappears as they master reading (Cornell, 1985). Another process that needs to be mastered is the association of each letter with its position in the word. Both developing and skilled readers can make errors characterised by internal letter transpositions (i.e., reading “from” as “form”) (Paterson et al., 2015). This effect is also very present in young readers and intensifies in the first years of learning, before diminishing to a minimum in expert readers (Grainger et al., 2012).

## Brain networks for reading

All the above-mentioned behavioural studies have made it possible to establish and confirm the presence of two distinct reading pathways enabling the transition from print to speech. Imaging studies, such as fMRI and MEG, have enabled us to gain a better understanding of the organisation of reading networks, and to study their development during learning.

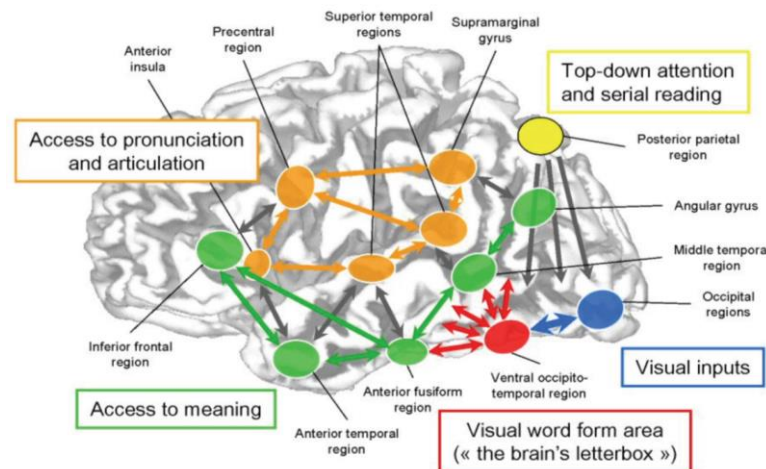
When we read, the visual system is the first to be called upon (Figure 4). This system, present in both hemispheres, is divided into different areas, some dedicated to recognising faces, others shapes, etc. The visual area specialised in recognising letters and written words is located in the left lateral occipitotemporal sulcus and is known as the Visual Word Form Area (VWFA) (Cohen et al., 2000). A central element in the processing of written words, it is found in all good readers and is specifically activated in the recognition of written words (Dehaene et al., 2002), regardless of the location of the word (presentation in the right or left visual field) (Cohen et al., 2000), of case (Dehaene et al., 2001) and of language (Bolger et al., 2005). It is also able to distinguish a word from its mirror image, unlike other visual areas (Pegado et al., 2011a).

The causal role of the VWFA in reading was highlighted by a surgical study of a patient suffering from epilepsy in the left occipitotemporal lobe. Removal of this lobe led to a deterioration in the patient's reading ability, transforming him from a fluent reader to a decoder with great difficulty (Gaillard et al., 2006). It is also in this region of the brain or its afferents and efferents that lesions are located in patients suffering from pure alexia (Cohen, 2003; Déjerine, 1892).

The results of functional imaging in patients with hemi alexia, who can only read information presented in the right visual field, demonstrate the central role of the corpus callosum in the transmission of information from the left visual field to the VWFA (Cohen et al., 2000). In these patients, the lesion of the corpus callosum breaks the invariance of VWFA activation, which is only activated for words presented in the right visual field, as the information received by the left visual field can no longer

circulate to the VWFA. The need for information from the left visual field to pass through the corpus callosum was also demonstrated by the behavioural results, which showed that the accuracy of processing of stimuli presented in the right visual field was slightly greater than that of the left visual field (Cohen et al., 2000).

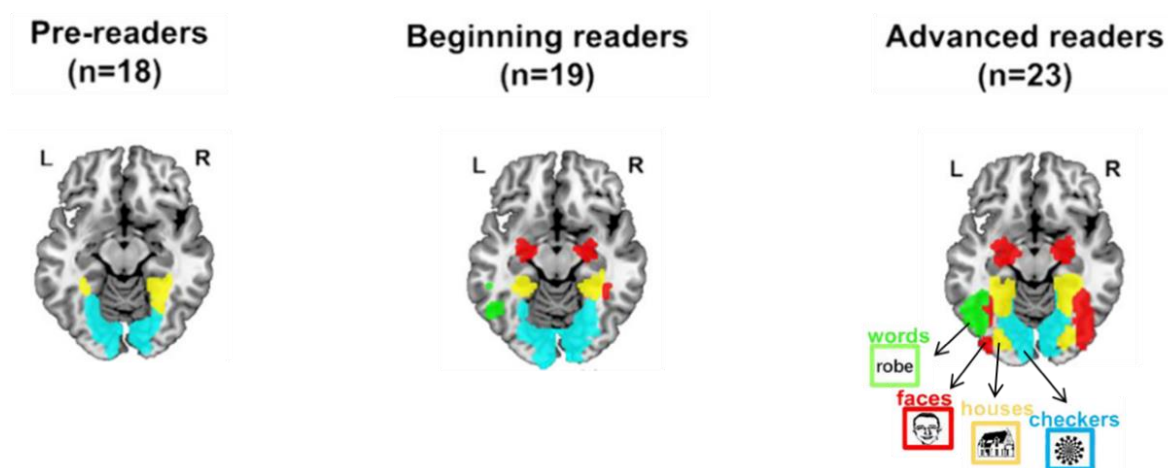
According to the neuron recycling hypothesis, the appearance of VWFA in adult readers results from the recycling of certain neurons in the left visual area (Dehaene et al., 2010). This study by Dehaene et al., which compares VWFA activation between illiterate and reading adults, also shows a strong correlation between reading ability and VWFA activation (Dehaene et al., 2010). Finally, the lateralization of the VWFA in the left hemisphere could be partly linked to its proximity to the language brain areas, which are also located in the left hemisphere in most right-handed adults (Monzalvo & Dehaene-Lambertz, 2013), and with whom the VWFA communicates in order to access the phonology and meaning of written words (Dehaene & Cohen, 2011).



**Figure 4.** A brain model for reading as described by Dehaene in its book *“The Reading Brain”*. The information first enters the visual system, where it is sent to the VWFA. Connections between the VWFA and the language areas then allow access to the pronunciation and meaning of the words read.

## How do these networks develop during reading acquisition?

Long before they learn to read, young children's language areas are already organised and activated in a similar way to those of adults (Dehaene-Lambertz et al., 2006). The brain area where the VWFA classically appears is also already connected to the language areas (Feng et al., 2022). Although the visual areas specialised in recognising shapes, houses and faces are also activated in a similar way to those of adults, the left lateral occipitotemporal sulcus is not yet activated specifically for words (Dehaene-Lambertz et al., 2018). The effect of the onset of reading in this area is visible from the very first months: the left lateral occipitotemporal sulcus, previously weakly specialised, specialises in letter and word recognition and the VWFA appears, in the same position as it does in adults, as depicted on Figure 5 (Dehaene-Lambertz et al., 2018; Feng et al., 2022).



**Figure 5.** Emerges from the VWFA during learning to read (Feng et al. 2022). It only takes a few months of learning to read for the first specific activations of the VWFA to appear, already in the same place as in normal adult readers.

The onset of VWFA is a consequence of learning to read and is not age-related. Before first grade, it is already present in children with some knowledge of reading (Feng et al., 2022; Monzalvo, 2011), and appears in nursery school children when they are encouraged to learn grapho-phoneme correspondences (Brem et al., 2010). Similarly, it is absent from the visual areas of illiterate adults, but appears when they learn to read (Braga et al., 2017; Dehaene et al., 2010).



These results support the hypothesis of neuronal recycling for learning to read and show that the cerebral plasticity of visual areas continues into adulthood. However, they show that it becomes less effective with age, with the pattern of activation during the first year of reading acquisition differing between ex-illiterates and young children. In young children, VWFA activation increases very rapidly, peaking after a few months of learning, indicating that reading is still very cognitively demanding. The gradual decrease observed over the following months marks the transition to automatic reading, which is less costly and more effective (Dehaene-Lambertz et al., 2018). In ex-illiterate adults, on the other hand, there was a constant increase in VWFA activation, while progress in reading was rather slow, a sign that reading is still a costly process, even after more than a year of learning (Braga et al., 2017). Finally, learning to read also affects language areas, significantly increasing the activity of areas associated with phonological representations and sentence integration after one year of learning to read (Monzalvo & Dehaene-Lambertz, 2013).

## **How does language affect these observations?**

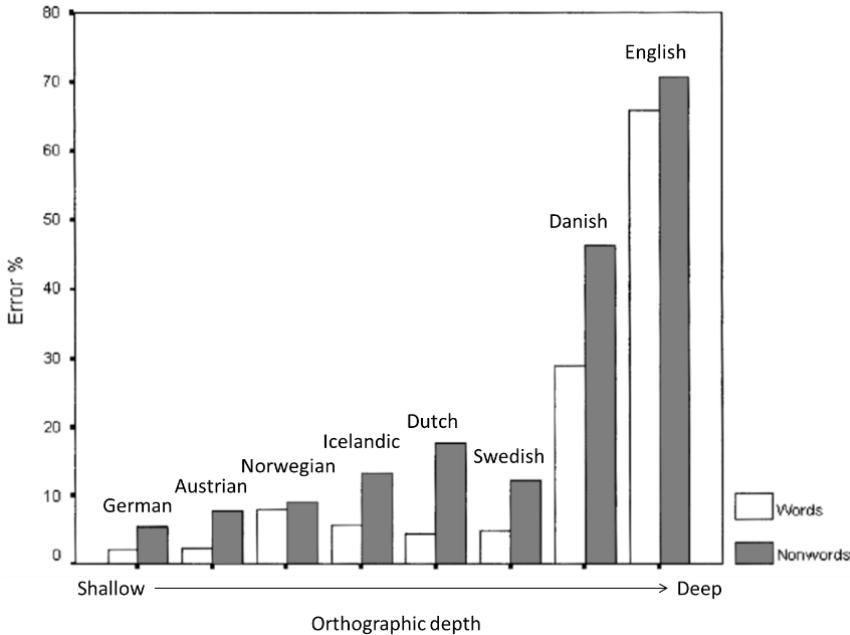
The previous paragraphs have shown that age, and therefore level of reading experience, affects both behavioural results and brain activity. Language is another variable that could have an impact on these observations. It would be illusory to attempt to summarise in a few lines the vast field of research into the differences between languages, but we will give you a brief overview of some of the major results, as these observations will be useful in contextualising the results of this thesis work within the existing literature.

Studying the differences between languages means first characterising the different writing systems according to the size of the encoded units (words, syllables, phonemes) as well as orthographic transparency, which characterises the degree of regularity between sounds and their orthographies, in both directions. In alphabetic languages, graphemes are associated with phonemes. If each grapheme is associated with a phoneme in a bijective relationship, the language is said to be

transparent (e.g., Finnish). As the number of irregularities increases, the language is considered increasingly opaque. By characterising the degree of irregularities in the grapheme-phoneme conversions for each language, we can see that Greek, Finnish, German, Spanish and Italian are transparent languages, while English is a very opaque language (i.e., a single grapheme may make many different phonemes, as in the case of the letter 'a' in words like 'cat', 'cane', 'caught' and 'America'). In the middle of this continuum are French, Portuguese and Danish (Ziegler et al., 2010). French, the language of the current work included many rule-based grapheme phoneme relations. For example, the letter 's' can make the /s/ or /z/ sound depending on if it is sandwiched between vowels or a consonant. There are also many frequent irregular words (i.e., the letter 'e' in the word for woman 'femme') and silent letters that act as morphological pointers (e.g., the third person plural silent 'ent' ending on a verb, which is an audible grapheme combination in non-verbs).

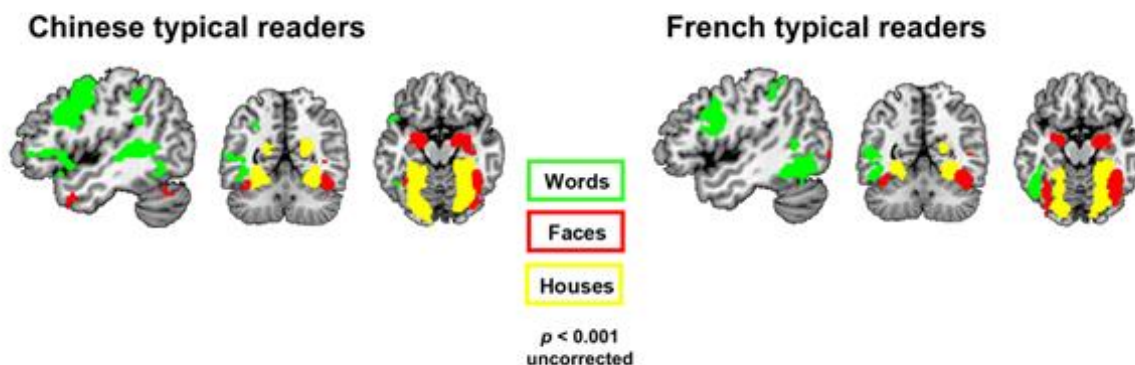
From a behavioural point of view, language opacity affects the speed of acquisition of fluent and efficient reading between languages. Several studies have highlighted this (Vaessen et al., 2010; Ziegler & Goswami, 2005). One of the most important observations of this difference was made when researchers from various European countries joined forces to create a database of words and non-words matched across 14 different languages. They had first-grade children from all countries read this pool of words, all of whom had received the same amount of reading instruction. In countries where spelling is transparent, pupils were at ceiling in both words and pseudowords reading by the middle of first grade. The success of French, Portuguese and Danish pupils was close to 75%, while that of English pupils, the most opaque language studied, did not exceed 40%, as shown on Figure 6 (Seymour et al., 2003). These results can be explained as follows. The more transparent a language is, the faster pupils acquire the rules of grapheme-phoneme conversion and can thus use their phonological route to decode words they are encountering for the first time. This acquisition becomes more laborious as the opacity of the language increases. It also becomes more difficult for readers to rely solely on these rules, as they can contain so many exceptions. Learning to read will then require learning to memorize

regular word lists and orthographic rules on the one hand or learning words through their common rhyme units, a process that necessarily takes longer to master (Ziegler & Goswami, 2005).



**Figure 6.** Influence of orthographic depth on the reading of simple words and nonwords in the middle of 1<sup>st</sup> grade (Seymour et al. 2003). German and Austrian-speaking children are already at ceiling when English-speaking children are still making 70% errors.

Despite these language differences, brain imaging studies show a universal reading circuit, unaffected by differences between languages (Figure 7). The reading networks, and in particular the VWFA, are present in the same way regardless of culture, in children and adults alike (Feng et al., 2020a; Nakamura et al., 2012; Rueckl et al., 2015). These results are in line with the neuronal recycling hypothesis for learning to read, which forces part of the visual system, identical in all languages, to specialise in letter recognition.



**Figure 7.** Brain imaging in Chinese and French typical children (Feng et al. 2020). Although Chinese and French are very different languages, brain activations exhibit a similar pattern and the VWFA is located in the same brain region.

## READING IN DYSLEXICS

Developmental dyslexia is a neurodevelopmental disorder characterised by important difficulties in reading acquisition in the presence of normal intelligence and access to education. According to the APA definition, it is a "neurologically based learning disability manifested as severe difficulties in reading, spelling, and writing words". Many different directions are pursued in dyslexia research, including the existence of subtypes, their behavioural characterization, and their cognitive, circuit-level, neuronal and genetic mechanisms. In terms of reading skills, dyslexic readers represents the tail of a normal distribution of word reading abilities (Shaywitz et al., 1992). In contrast to acquired dyslexia, which results from a neural lesion to a fully developed system, developmental dyslexia is a disorder that prevents the developing reading system from becoming efficient and automatized. In this thesis work, we only focus on developmental dyslexia.

### **How do behavioural results and brain activity vary with dyslexia?**

In the first part of this introduction, we focused on the impact of different variables such as length, frequency and lexicality on reading in adults and normal children and on the brain networks involved

in reading. To understand how these effects vary with dyslexia, numerous behavioural studies comparing dyslexics and non-dyslexics have been carried out. These studies use two types of control group: chronological age controls and reading level controls. Reading level controls are important to ensure that the differences observed between dyslexics and chronological age controls are not merely a manifestation of developmental delay. On overall performance in different reading tasks, most studies converge on the following observations: dyslexic pupils are slower and make more errors than aged matched controls (Araújo et al., 2014a; Barca et al., 2006; Bergmann & Wimmer, 2008; Davies et al., 2013; Di Filippo et al., 2006a; Martens & De Jong, 2006; Ziegler et al., 2003; Zoccolotti et al., 2005a) but they make just as many errors and are just as fast as reading level matched controls (Figure 8) (Davies et al., 2013; Martens & De Jong, 2006; Ziegler et al., 2003; Zoccolotti et al., 2005a).

The lexicality effect, which indicates a more rapid and precise response to words than to pseudowords, is more marked in dyslexic pupils than in aged-matched controls (Araújo et al., 2014a; Bergmann & Wimmer, 2008; Martens & De Jong, 2006; Ziegler et al., 2003). However, this difference in raw scores disappears when we check for the overall slowness of the dyslexic children (Di Filippo et al., 2006a). Comparison with reading level controls produces more controversial results. While differences between these two groups are observable, some studies report a greater lexicality effect in dyslexics than in reading level control (Martens & De Jong, 2006; Ziegler et al., 2003) while others achieve the opposite results (Davies et al., 2013).

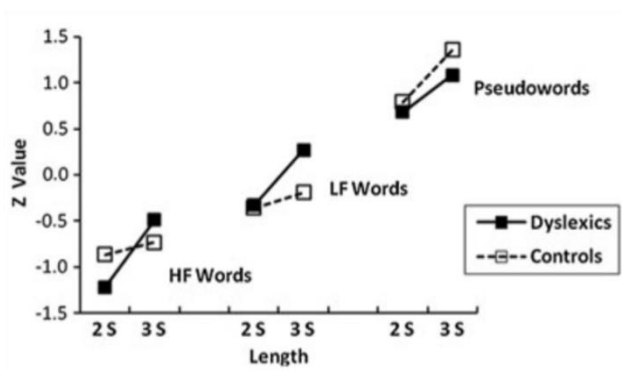
Compared to chronological age controls, dyslexics show a massive length effect on both RT and accuracy (Davies et al., 2013; De Luca et al., 2008; Martens & De Jong, 2006; Spinelli et al., 2005; Ziegler et al., 2003), including for words (Araújo et al., 2014a; Di Filippo et al., 2006a; Zoccolotti et al., 2005a). The linear relationship between length and response time is very similar to that for reading level controls (Martens & De Jong, 2006; Zoccolotti et al., 2005a), and sometimes even higher in dyslexics (Ziegler et al., 2003). Contrary to the findings on the lexicality effect, once controlled for the overall slowness of dyslexic participants, this length effect still exists and is still higher for dyslexics than for

age-matched controls (Di Filippo et al., 2006a), which accentuates the real existence of such a difference.

The results reported on the frequency effect are a bit less consistent between studies. While some studies report a greater frequency effect in dyslexics than in age-matched controls (Araújo et al., 2014a; Barca et al., 2006), others report similar effects (De Luca et al., 2008). Comparison with reading level controls shows that the frequency effect is greater in dyslexics (Beech & Awaida, 1992). Finally, with regard to the regularity effect, several studies have shown no difference between the results observed in dyslexic children and those observed in controls of the same age or reading level (Marinus & De Jong, 2010; Ziegler et al., 2003).

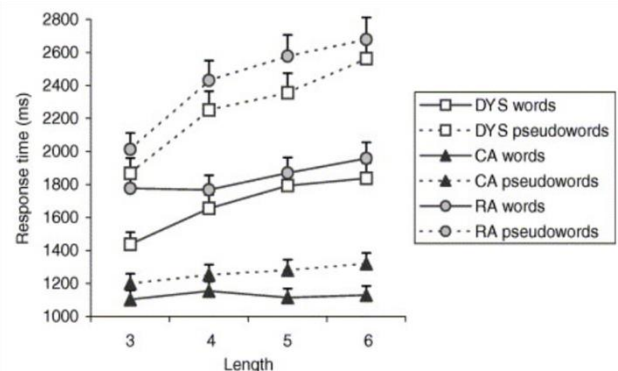
From the dual-route perspective, dyslexic individuals exhibit reading profiles that bear resemblance to those of younger children. Their pronounced length effect, even when reading regular words, highlights their heavy reliance on sublexical processes, underscoring that their reading is not yet automatic. Additionally, the presence of a frequency effect suggests that dyslexic individuals, when encountering familiar words, can also tap into their lexical route for reading.

a. Interaction between lexicality, frequency and length in dyslexic children and chronological-aged controls



(Araújo et al., 2014)

b. Interaction between lexicality and length in dyslexic children and both chronological-aged (CA) and reading-level (RA) controls



(Martens & De Jong, 2006)

**Figure 8.** Illustration of the differences in the effects of lexicality, length and frequency and their interactions between dyslexic children and typical readers. **(a) Interaction between lexicality, frequency and length in**

**dyslexic children and chronological-aged controls.** Dyslexic children exhibit a strong length effect on both high frequency and low-frequency words, conversely to chronological-aged controls (Araújo et al., 2014). **(b) Interaction between lexicality and length in dyslexic children and both chronological-aged (CA) and reading-level (RA) controls.** Dyslexic children reading profile is very similar to that of reading-level controls but differs from that of chronological-aged controls: chronological-aged controls are faster and they exhibit no length effect for words and a small length effect for pseudowords, whereas dyslexics and reading-level controls shows strong length effect for both words and pseudo-words (Martens & De Jong, 2006).

Other differences have been observed between the brains of normal readers and those of dyslexics. While the results of many studies agree that brain volume is smaller in dyslexics than in normal readers (Ramus et al., 2018), the differences in terms of brain activity are more open to debate. Some studies have shown significant differences in activation in certain areas involved in reading and language (Altarelli et al., 2014; Feng et al., 2020a), a recent review by Ramus et al. (2018) shows, however, that care must be taken when interpreting these differences, which depend very much on sample size and the statistical analysis methods used. When we looked at the differences between languages, we observed that all these observations, both behavioural and in terms of imagery, are only marginally modulated by orthographic depth (Feng et al., 2020a; Ziegler et al., 2003).

## **Early categorization of dyslexia profiles**

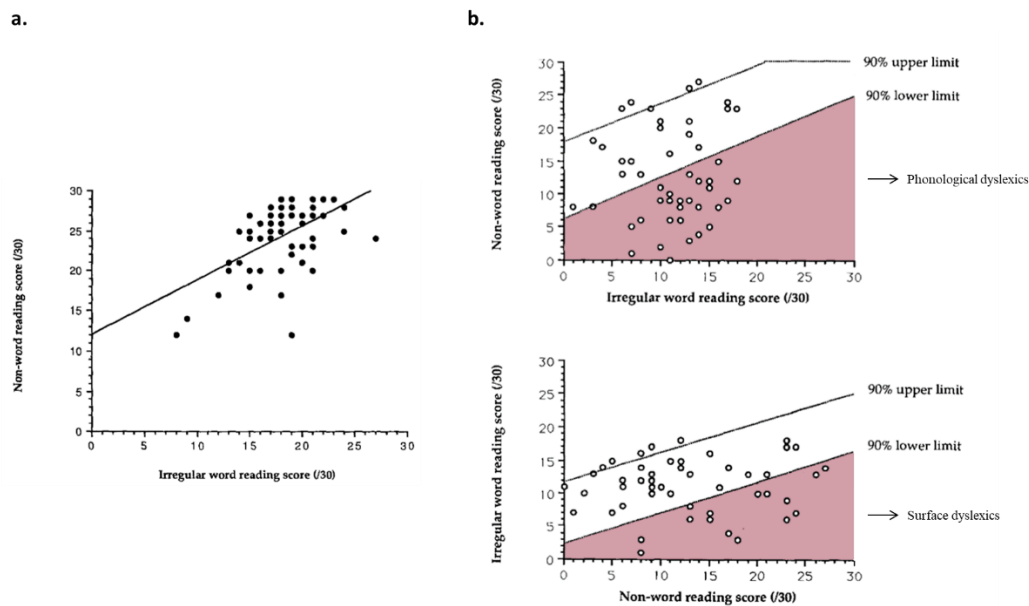
Critically, these studies consider overall errors and reading response times and not the difference in the types of errors that dyslexia may produce compared to normal reading. Nor do they consider that dyslexia may arrive from a developmental deficit at any stage in reading, thus producing as many different types of dyslexia profiles as there are processing components in the brain to reading. If there are differences between dyslexics and normal readers, there are also differences between dyslexics themselves. Numerous studies, on single cases or small cohorts, describe a wide variety of error profile between their participants. Some dyslexics have great difficulty in reading irregular words (Castles & Coltheart, 1993; Manis et al., 1996; Sprenger-Charolles et al., 2000, 2011; Stanovich et al., 1997; Valdois et al., 2003), others in reading pseudowords (Castles & Coltheart, 1993; Manis et al., 1996; Sprenger-Charolles et al., 2000, 2011; Stanovich et al., 1997; Valdois et al., 2003), others confuse mirror letters (b, d, p and q, u and n, m and w) (McCloskey et al., 1995; McCloskey & Rapp, 2000b), others

make a lot of transposition errors (Friedmann & Haddad-Hanna, 2012; Friedmann & Rahamim, 2007a; Güven & Friedmann, 2019; Kohnen et al., 2012; Potier Watkins et al., 2023), and still others swap the letters of a word with those of surrounding words (Friedmann, Kerbel, et al., 2010; Friedmann & Haddad-Hanna, 2014; Potier Watkins et al., 2023; Rayner et al., 1989).

In the early 1990s, these findings led to the development of the first screening tools for categorizing dyslexia subtypes based on the dual-route model of reading. Three primary subtypes emerged: surface dyslexia, phonological dyslexia, and mixed dyslexia. Surface Dyslexia was characterized by a deficit in the lexical route and is typically identified through assessments involving irregular word reading. Individuals with surface dyslexia heavily rely on sublexical reading, resulting in slower reading and a tendency to apply regular phonetic rules to irregular words, which makes reading irregular words challenging. Phonological dyslexia was associated with a deficit in the sublexical route. In this case, individuals struggle with phonological decoding and must rely on memorization to read words, slowing down their reading progress, as they cannot easily sound out words. Phonological dyslexia is often diagnosed by assessing the individual's difficulty in reading pseudowords that does not exist. Mixed dyslexia can encompass deficits in both the lexical and sublexical routes. Individuals with mixed dyslexia encounter challenges in recognizing familiar words and decoding unfamiliar words. Their reading difficulties span a broad range of words, combining characteristics of both surface and phonological dyslexia. Numerous studies have examined and described these dyslexia profiles, but these definitions remain limited given the complex nature of reading. Using dual-route model, researchers have aimed to provide a framework for understanding these distinct reading profiles and tailoring interventions to address the many different profiles of dyslexia observed in the literature. (Castles & Coltheart, 1993; Manis et al., 1996; Sprenger-Charolles et al., 2000, 2011; Stanovich et al., 1997). In their 1993's study, Castles et Coltheart also suggested to characterize soft types of phonological and surface dyslexia, dyslexic people impaired in both procedures, but much more in one than in the other, using regressions, as explained in Figure 9 (Castles & Coltheart, 1993). Most dyslexics would therefore be soft types. Castles and Coltheart's regression technique was subsequently reused



in other studies which gave fairly similar results in terms of the prevalence of the different subtypes of dyslexia, with a higher proportion of phonological dyslexics in all but one of these studies (Castles & Coltheart, 1993; Manis et al., 1996; Sprenger-Charolles et al., 2000, 2011; Stanovich et al., 1997; Ziegler et al., 2008). The variations observed can be attributed both to the age of the participants (younger in Stanovich et al., for example) and to the language.



**Figure 9.** The regression technique proposed by Castles and Coltheart to identify phonological and surface dyslexics (Castles & Coltheart, 1993). **(a)** Using age-matched controls, they first extract the linear relationship between the number of irregular words and pseudowords correctly read. They used this regression to find the boundaries below which student should fall to conclude that their performance was significantly lower than expected. **(b)** The scatterplots of dyslexics' results using these boundaries showed the existence of 16 surface dyslexics and 29 phonological dyslexics.

## Theories on the cognitive basis of dyslexia

Samuel Orton was one of the first to report dyslexia cases and tried to find a cause to that disorder. According to him, this deficit, which he named "strephosymbolia", was a brain-based disorder. For him, the visual form of letters was correctly stored in the left hemisphere but mirrored in the right hemisphere. In normal readers, the left hemisphere was dominant, so only the correct image emerged,

unlike dyslexic readers, for whom there was an incomplete suppression of the mirror image, leading to the huge number of letter confusions that he observed (Orton, 1925, 1928).

This hypothesis, which is no longer taken seriously today, nevertheless helped to lay the foundations for dyslexia and set in motion the thinking that gave rise to the various theories presented below.

### ***Various theories of a core deficit in dyslexia***

#### *Visual theories*

Since reading involves the association of letters and sounds, it initially requires a good perception of words. This is why several studies have focused on visual deficits in dyslexics. One of the recent hypothesis is the theory of Le Floch and Ropars, according to which an anomaly in Maxwell's spot centroids, in the central region of the fovea, is the cause of all dyslexias (Le Floch & Ropars, 2017). Dissymmetrical in normal readers, this task would be symmetrical in all dyslexics, leading to the formation of mirror images and consequent confusion of letters, particularly the mirror pairs b/d and p/q. For them, this would explain why dyslexics struggle with reading. A retinal anomaly, if it was properly documented, would however not explain the dissociations observed in dyslexia, for instance between number and letter reading (Dotan & Friedmann, 2019; Friedmann, Dotan, et al., 2010) and to the best of our knowledge, the results of this study have never been replicated.

Another hypothesis, this time better documented in the literature, makes a link between dyslexia and dysfunction of the magnocellular pathway, one of the two input pathways of the visual system (Boden & Giaschi, 2007; Gori & Facoetti, 2014; Livingstone et al., 1991). In their review, Stein & Walsh clearly show that there is a link between dyslexia and difficulties in processing movement or stimuli with low contrast, high temporal frequency or low spatial frequency, skills linked to the functioning of the magnocellular pathway. They suggest that these physiological differences between dyslexics and normal readers are at the root of reading problems in dyslexics (Stein & Walsh, 1997). Vidyasagar and

Pammer go further and suggest that a deficit in the magnocellular pathway is at the root of dyslexia because it interacts with deficits in visual attention (Vidyasagar & Pammer, 2010), defined as difficulties in shifting attention and orienting attention in the visual modality. The claim that dyslexia unilaterally stems from a visual-attention deficit has first been suggested by Vidyasagar in 1999 (Vidyasagar, 1999), even though these disorders are always associated with phonological disorders (Valdois, 2014). According to this hypothesis, the visual attention difficulties observed in dyslexics would interfere with the scanning of the sequence to be read, resulting in a defect in grapheme segmentation and, consequently, a defect in phonological awareness, which is known to be better when the child is aware of grapheme segmentation.

### *Phonological theory*

The phonological theory is the other major theory opposing the various visual theories mentioned above (Ramus, 2001, 2003a; Vellutino et al., 2004). In its strong version, it explains that dyslexia is specifically linked to a phonological deficit, characterised by a deficit in the ability to perceive and manipulate sounds. In particular, this deficit impedes the learning of the rules of grapho-phonemic conversion, which is essential for the acquisition of reading, as presented in the first part of this introduction. This theory is based in particular on the very poor performance of many dyslexics in phonological awareness exercises (Bradley & Bryant, 1983; Elbro & Jensen, 2005; Landerl et al., 2013; Sprenger-Charolles et al., 2000; Vellutino et al., 2004). Studies comparing the phonological deficit theory with others have shown that most dyslexics have a phonological deficit, and they postulate that their dyslexia is aggravated by the presence of visual or motor deficits, suggesting that a phonological deficit is inherent in dyslexia (Ramus, 2003b; Ramus et al., 2003; Saksida et al., 2016; White et al., 2006; Ziegler et al., 2008). The theory of the phonological whole thus explains the visual disorders described above as a consequence of dyslexia and not as a cause (Goswami, 2015a). It also seeks to explain the different profiles of dyslexia, whether surface, phonological or mixed. In the studies of different hard and soft types of dyslexia presented above which evaluate phonological skills, results showed that all

dyslexics, whether they present a profile of surface dyslexia or phonological dyslexia, were impaired (Manis et al., 1996; Sprenger-Charolles et al., 2000; Ziegler et al., 2008). Thanks to their longitudinal data, Sprenger-Charolles and colleagues were also able to trace pupils' performance back to the very beginning of learning to read, and they showed that even if some pupils went on to display different dyslexia profiles, they were all very similar at the outset (Sprenger-Charolles et al., 2000). They conclude that the two forms of dyslexia do not therefore have fundamentally different cognitive profiles, but rather are linked to different compensatory strategies, probably explained by environmental factors. A final surprising effect common to all these studies, is the virtual disappearance of surface dyslexics when the regression technique is applied in comparison with reading-level controls (Manis et al., 1996; Sprenger-Charolles et al., 2000, 2011; Stanovich et al., 1997). For the authors of these studies, phonological dyslexia is therefore the result of a genuine phonological deficit, while surface dyslexia is more likely to be linked to an overall delay in learning to read.

However, while the importance of phonological disorders in some dyslexics is no longer in doubt, considering phonological disorders as the sole cause of dyslexia is widely questioned, particularly in view of the existence of a number of cases of dyslexics with no phonological disorders (Castles & Friedmann, 2014; Valdois et al., 2003, 2011; Zoccolotti & Friedmann, 2010). Thus, other theories have emerged, encompassing the phonological deficit in a broader model.

### ***Dyslexia resulting from multiple deficits***

#### *Phonological deficit or visual-attention span deficit*

In Valdois' theory, dyslexic disorders are caused by two major deficits: a phonological deficit and a deficit of the visual-attention span, defined as the quantity of distinct visual elements that can be processed in parallel in a multi-element network. First introduced in 2007, the concept of a visual-attention span deficit in some dyslexics has since been the subject of many publications. While single case studies have demonstrated the existence of such a disorder and its dissociation from phonological

disorders (Valdois et al., 2003, 2011), cohort studies have shown that this type of deficit is fairly frequent and again, dissociated from phonological disorders (Bosse et al., 2007; Germano et al., 2014; Valdois et al., 2004; Zoubrinetzky et al., 2016). According to Valdois and colleagues, this deficit should not be considered marginal, since it concerns 18% of the population studied in a cohort of over a hundred children (Valdois et al., 2021). This type of disorder has already been identified in several languages of different opacity (Bosse et al., 2007; Germano et al., 2014), and the evidence that a proportion of dyslexics suffer from a visual-attentional span disorder is fairly solid. It is the purely visual origin of this deficit, and its causal relationship with dyslexia, that are still being questioned today, particularly by supporters of the phonological theory (Banfi et al., 2018; Collis et al., 2013; Goswami, 2015a, 2015b; Yeari et al., 2017; Ziegler et al., 2009). Several studies and reviews by Valdois emphasise the purely visual origin of visual-attention span disorder, which affects a wide variety of stimuli, and not just alphabetic stimuli (Lobier et al., 2012; Valdois, 2022; Valdois et al., 2012). Valdois and colleagues underline the causal nature of the disorder by studies comparing dyslexics and reading level controls, which have shown that the reduction in visual-attention span remains greater in dyslexics (Bosse & Valdois, 2003; Zoubrinetzky et al., 2014), as well as by remediation studies, in which training of the visual-attention span led to an increase in the reading ability of the children concerned (Valdois et al., 2014; Zoubrinetzky et al., 2019). At the present time, however, the lack of large-scale longitudinal studies remains a factor opposing this causal relationship, preventing the debate from being concluded (Goswami, 2015b).

With this theory, Valdois finally calls into question the surface/phonological/mixed categorisation of dyslexia subtypes established using the word list reading method of Castles and Coltheart (1993). While the method indicates that the vast majority of phonological dyslexia hard cases suffer from phonological disorders and that the vast majority of surface dyslexia hard cases suffer from a deficit in the visual-attention span (Valdois et al., 2003; Zoubrinetzky et al., 2019), cases of mixed dyslexia, although grouped under the same name, present very varied profiles. A study by Zoubrinetzky et al. of a group of mixed dyslexics found that a number of dyslexics had an isolated phonological disorder,

others had an isolated visual-attention deficit, others had both disorders and still others had neither. Finally, this study no longer looked at the error rate for each type of word, but rather at the error itself, which could be the result of omission, voicing confusion, digram parsing errors or contextual errors for instance. The study concludes that dyslexics with a phonological disorder will make more voicing errors, while dyslexics with a visual-attention span deficit will make more parsing errors (Zoubrinetzky et al., 2014).

#### *Multiple dyslexias arising from any site within the dual-route model*

Examining the specific errors made by a child is a method recommended by Friedmann and colleagues to distinguish between different dyslexia profiles and gain insights into potential underlying deficits. She further elaborated this approach by using stimuli that can elicit types of errors. For instance, if a reader has a deficit in recognizing letter positions, presenting anagrams should result in a higher rate of error production compared to readers with other types of dyslexia. This is because anagrams provide an opportunity for errors to occur without immediate self-correction, allowing the reader to produce an incorrect word also stored in their mental lexicon (for instance, reading 'silent' as 'listen'). Her model, based on the dual route model of reading, stipulates that a certain subtype of dyslexia appear as soon as one of the stages in the dual route model is impaired, giving a more precise categorization than the one first mentioned by Castles and Coltheart (1993). In addition to phonological and surface dyslexias resulting from impairment of the lexical or sublexical routes, a certain number of dyslexias can thus emerge from impairment at the orthographic-visual analysis stage. This is the case with letter position dyslexia, which selectively impairs the ability to encode the relative position of letters within words. Letter migrations occur most often between two adjacent internal letters, when the error creates another existing word (slime → smile or form → from), both when reading aloud and silently. Also observed in letter position dyslexia are errors of letter doubling (i.e., reading 'bile' as 'bible') or omission of one instance of a letter that appears twice in a word (i.e., the opposite, reading 'bible' as 'bile') (Friedmann & Rahamim, 2007b). Cases of pure letter-position

dyslexia, with no phonological deficit, have been demonstrated in a large number of languages, with a wide range of orthographic depths (Friedmann & Haddad-Hanna, 2012; Friedmann & Rahamim, 2007a; Güven & Friedmann, 2019; Kohlen et al., 2012; Potier Watkins et al., 2023). The observed dissociation between letters and numbers assessed that letter position dyslexia is domain-specific (Friedmann, Dotan, et al., 2010; Güven & Friedmann, 2019).

Attentional dyslexia is another orthographic-visual dyslexia which selectively alters letter to word binding. Here, letters are correctly identified and almost always retain their correct relative position, but they migrate towards a neighbouring word (read fig-tree as fig-free). Migration can occur both horizontally and vertically, most often at the end of a word, and more frequently from the first to the second word in Hebrew (Friedmann, Kerbel, et al., 2010) and from the second to the first word in French (Potier Watkins et al., 2023). It takes the form of substitutions, additions or omissions of letters in neighbouring words (Friedmann, Kerbel, et al., 2010; Friedmann & Haddad-Hanna, 2014; Potier Watkins et al., 2023; Rayner et al., 1989). Although it might be thought that this type of dyslexia reflects a deficit in visual attention, a study by Lukov et al. shows that these two deficits can be uncorrelated (Lukov et al., 2015). In a recent study, Toledano and Friedmann point out that between-words migrations, when reading aloud, is also characteristic of dyslexics with a phonological output buffer deficit (Toledano & Friedmann, 2023). Silent reading and the repetition of word pairs enable these two types of disorder to be differentiated: attentional dyslexics will also make errors in silent reading but not when repeating word pairs, unlike dyslexics with a phonological output buffer deficit. Other rarer types of dyslexia also have their origins in visual-orthographic analysis. In particular, there may be a deficit in encoding the abstract identity of letters. This is a rare form of dyslexia, in which patients are unable to read because they are unable to identify letters correctly. They cannot name a letter, identify a written letter by its name or sound, or match upper and lower case letters (Brunsdon et al., 2006). Finally, a deficit in the output of the orthographic visual analyser that affects all functions of the orthographic-visual analyser leads to visual dyslexia, a profile in which participant omit, substitute or add letters in words, no matter the letter and no matter their position (Friedmann et al., 2012).

Friedmann's work has also made it possible to refine the different types of surface dyslexia that exist. The lexical route is made up of different stages, each of which can be altered. Lukov and Friedmann identified three subtypes of surface dyslexia, but in reading aloud all these variants show similar patterns (Friedmann & Lukov, 2008), with a high difficulty in reading irregular words.

The existence of these different subtypes of dyslexia, in their pure form, corroborates that of each of the stages of the dual route model of reading as described by Friedmann and Coltheart, as well as their independence (Friedmann & Coltheart, 2018). These different deficits may also coexist in the same individual. At present, however, there is no precise measure of the prevalence of each of these types of dyslexia in the population.

## **Dyslexia remediation**

Seeking to help dyslexic patients is usually the first step after diagnosis. A vast field of research has looked at the most effective methods. Some research has looked at the impact of cognitive re-education, training fundamental aspects of reading such as decoding, phonological awareness or visual-attention span. Several studies have shown the effectiveness of intensive work on one or the other, depending on dyslexia profile, with effects that are still visible in the long term (Alexander et al., 1991; Elbro & Jensen, 2005; Heikkilä et al., 2013; Saine et al., 2011; Temple et al., 2003; Torgesen et al., 2001; Valdois et al., 2014; Zoubrinetzky et al., 2019). These intensive training sessions are given by speech therapists, or using software such as Play-On®, RapDys®, GraphoGame® or MAEVA®.

Although these re-educations are very effective, they nevertheless require considerable human resources and a great deal of patience (intensive training of 67.5 hours spread over 8 weeks for each child, individually with an experienced educational therapist and a follow-up of weekly 50-minute sessions for the rest of the year for the Torgesen et al. study / 100 minutes a day, 5 days a week for 5 weeks for the Temple et al. study / 64 hours or training spread over 6 weeks for the Alexander et al. study / 30 minutes a day for 17 weeks for the Elbro and Petersen study / 15 minutes a day during 12



weeks for the Zoubrinetzky et al. study / 6 days a week during 6 weeks for the Valdois et al. study). This is why another field of research has focused on facilitating reading for dyslexics by manipulating their reading experience.

Increasing font size and spacing is one easier to put in place aids that has proved effective so far, with reading made easier for dyslexics using a larger font size (O'Brien et al., 2005; Rello & Baeza-Yates, 2017). While it may simply indicate that many dyslexics, similar to beginner readers, have not yet adapted to small print, it does offer a simple way to help them. Increasing the spacing between characters is also an effective parameter (Zorzi et al., 2012), that could even be more effective than increasing the font size (Katzir et al., 2013). Spacing letters a few extra percent apart has been reported to increase reading speed, reduce errors and facilitate comprehension (Łuniewska et al., 2022; Stagg & Kiss, 2021; Duranovic et al., 2018; Rello & Baeza-Yates, 2017; Hakvoort et al., 2017; Sjoblom et al., 2016; Zorzi et al., 2012). To be optimal, it has to be combined with an increase in the spacing between words (Galliussi et al., 2020). Other parameters such as coloured background, the use of specific fonts or line spacing have also been studied, with no facilitating effect found (Rello & Baeza-Yates, 2017). The effectiveness of specific fonts which is occasionally reported (Bachmann & Mengheri, 2018) fades once controlling for the spacing between characters, which is greater in specific fonts (Joseph & Powell, 2022; Galliussi et al., 2020; Marinus et al., 2016), thus suggesting that character spacing, rather than the font itself, is the most impactful variable. For easier reading, some data also suggest that the style of the font should be sans serif and should avoid italics (Rello & Baeza-Yates, 2013). Finally, single cases and group studies reported that colouring each letter using a different colour did not improve their dyslexic participants' results, and in some cases even worsen them, compared to baseline (Friedmann & Rahamim, 2014; Humphreys & Mayall, 2001; Koornneef & Kraal, 2022; Pinna & Deiana, 2018).

Regarding specific subtypes of dyslexia, such as attentional or letter position dyslexia, Friedman et al. have found that the use of an attentional window drastically reduced errors in attentional dyslexics (Friedmann, Kerbel, et al., 2010), while following words with the finger greatly helped patients with

letter-position dyslexia (Friedmann & Rahamim, 2014). Other studies have shown that intensive training, based on recognising letters and their sounds, enabled a child with a letter-processing impairment to make progress in his ability to associate graphemes and phonemes, a fundamental stage in beginning to read (Brunsdon et al., 2006).

## **GOAL OF THE THESIS: NEW TOOLS FOR DIAGNOSING AND REMEDYING DYSLEXIA**

Building on the theoretical background described above, this thesis aims to develop and evaluate new methods for the diagnosis and remediation of dyslexia. The CERENE schools welcome many pupils at the start of 6<sup>th</sup> grade. In **chapter one** of this work, we will therefore present the lexical decision tool that we have developed and normed on 1501 6<sup>th</sup> graders, with the aim of allowing teachers at CERENE and elsewhere to measure pupils' reading difficulties in terms of lexicality, length, and frequency effects, at a pivotal age when reading plays an increasingly important role in academic success. This test also made it possible to study, from a more theoretical point of view, the impact of reading fluency level on lexical decision results in years-of-education equivalent non-dyslexic students.

It is important to be able to assess reading accurately and quickly at the start of 6<sup>th</sup> grade, but it is also essential to be able to detect specific deficits in younger children. Reading tests used by practitioners and the scientific community alike to quantify reading fluency, only provide quantitative measures of overall reading speed and accuracy, and do not look at the exact errors made by the child. To provide CERENE with a more informative, but equally effective test, in the **chapter two** I present the Mariette, a nonsense text in which we characterized the exact errors of 812 pupils from 1<sup>st</sup> to 5<sup>th</sup> grade. Error thresholds beyond which a child is likely to present a specific deficit were established for dyslexias previously reported in the literature. The test also quantifies the child's fatigability and word skipping.

It is recommended that children with dyslexia receive 1 hour of extra individualized care from a speech therapist. Finding the personnel trained to meet this demand, not to mention the required for the child and parent to manage this extra-curricular goal, is becoming harder and harder. More and more parents are therefore looking for alternative solutions. Unfortunately, many of these quick fixes such as Lexilight<sup>®</sup> lamp and Lexilens<sup>®</sup> glasses are untested, although claiming that their technique using light flickering, alleviates difficulties in reading for up to 90% of dyslexics. CERENE has reported being regularly asked by parents whether buying these products could help their children. Given the lack of scientific evidence on these devices, [chapter three](#) presents the study we conducted on the impact of high- and low-frequency flickering on the reading performance of 375 normal adult readers and 22 dyslexic children. The results we obtained are in stark contrast with marketing claims and emphasize the role of rigorous scientific research in empowering dyslexic individuals to make informed decisions.

All these results have been established with the goal of practical applications. This is why all these chapters will also be published in French. The lexical decision test developed for 6<sup>th</sup> graders is also freely available online, and the Mariette will be made available soon once all our analyses have been completed.

**CHAPTER 1 – EXAMINING THE IMPACT OF READING  
FLUENCY ON LEXICAL DECISION RESULTS ON FRENCH  
6TH GRADERS**

## **ABSTRACT**

Purpose: How do lexical decision results vary in students with the same amount of education (all students were in their first year of middle-school), but different levels of reading fluency? Here, we tested a prediction of the dual-route model: as fluency increases, variations in the results may reflect a decreasing reliance on decoding and an increasing reliance on the lexical route.

Method: 1,501 French 6<sup>th</sup> graders passed a one-minute speeded reading-aloud task evaluating fluency, and a ten-minute computerized lexical decision task evaluating the impact of lexicality, length, word frequency and pseudoword type.

Results: As predicted, the word length effect varied dramatically with reading fluency, with the least fluent students showing a length effect even for frequent words. The frequency effect also varied, but solely in proportion to overall reading speed, suggesting that frequency affects the decision stage in all readers, while length impacts poor readers disproportionately. Response times and errors were also affected by pseudoword type (e.g., letter substitutions or transpositions), but these effects showed minimal variation with fluency.

Conclusion: Our results highlight the variability in middle-school reading ability and describe how a simple lexical decision task can be used to assess students' mental lexicon (vocabulary) and the automatization of reading skills.

## **INTRODUCTION**

Lexical Decision (LD) is a classic psycholinguistics paradigm requiring participants to classify visually presented stimuli as words or pseudowords. It has been widely used in cognitive science and provides well-replicated measures of key components of the visual word recognition process, which are captured by the Dual Route Model of reading (Coltheart et al., 2001b; Fiebach et al., 2002). This model describes the reading process as the outcome of two competing pathways. On the one hand, the lexical pathway lends itself to fast fluent reading of known words stored in the mental lexicon. This lexicon is affected by word frequency, such that most frequent words are quicker to retrieve than less frequent ones. On the other hand, the sublexical pathway enables reading unknown words or pseudowords, via the application of grapheme-phoneme correspondence (GPC) rules. The necessity to decode in order to access sound and thereby meaning (or rejection of meaning) causes response time (RT) to stimuli processed via the sublexical pathway to be strongly affected by the number of letters in the word (Acha & Perea, 2008; Bijeljac-Babic et al., 2004b; Di Filippo et al., 2006b; Martens & de Jong, 2006).

With LD, it is very easy to measure the length, lexicality and frequency effects that arise from the use of one of these two routes by varying the lexicality of stimuli (word or pseudoword), their length or their frequency (for words only). Numerous longitudinal studies have investigated how these effects change with reading acquisition. For example, Children, compared to adults, are affected by a more pronounced word length effect, as they rely more heavily on sublexical procedures when learning to read (Acha & Perea, 2008). As reading becomes efficient, the word length effect fades, indicating that the lexical pathway becomes fully operational, meaning that reading has become automatic (B. Weekes, 1997; Zoccolotti et al., 2005b). This decrease arises between 3<sup>rd</sup> and 5<sup>th</sup> grade, indicating a slow transition from serial grapheme–phoneme mapping to a greater reliance on lexical knowledge (Bijeljac-Babic et al., 2004b).

The decrease in the length effect is concomitant with the appearance of a frequency effect (Faust, Balota, Spieler, Ferraro, et al., 1999). The improvement in word reading performance is observed with longer words, plateauing within a range of approximately 3-8 letters (New et al., 2006b). Conversely, response time (RT) and accuracy are increasingly influenced by word frequency (Brysbaert et al., 2011; Burani et al., 2002; Grainger & Segui, 1990; Ratcliff et al., 2004). At this stage, reading is characterized by a linear relationship between the logarithm of frequency and RT (Norris, 2006). This maturing of faster lexical procedures can take from one to two years, depending on the transparency of the orthography in a given language (Schmalz et al., 2013b). In this study, we focus on the French language, which has an opaque orthography characterized by numerous graphemes with diverse phonemic realizations. It is noteworthy that the frequency effect in reading acquisition typically appears after nearly a year (Sprenger-Charolles et al., 1998). Additionally, the lexicality effect generally emerges for high-frequency words around the 3<sup>rd</sup> grade (Araújo et al., 2014b; Di Filippo et al., 2006b; Juphard et al., 2004b; Sela et al., 2014).

Looking at the results of the studies mentioned above, one could naively imagine that within the same grade, the effects observed differ little between students since they have all been exposed to reading for the same number of years. However, it is well known that large differences in reading literacy exist within grade. This is particularly true at the start of middle-school in France (6<sup>th</sup> grade), where this study took place, where many students enter without the skills needed for independent reading (Andreu et al., 2021, 2022). To our knowledge, the only current studies that look at students in the same grade compare dyslexics and non-dyslexics (Castles, 2006; Martens & de Jong, 2006; Zoccolotti et al., 2005b). Longitudinal studies are limited to a few hundred students per grade, which does not provide sufficient statistical power to study the effects of student's reading level within each grade. In their recent study, Yeatman et al. looked at the impact of reading level on length effect only, but through a wide range of ages (Yeatman et al., 2021). With our large-scale study conducted on a representative sample of the French 6<sup>th</sup> graders, we aim to better understand how do lexical decision

results vary within a grade, and how answer to pseudowords is affected by the reading level of the student.

### **How do lexicality, length and frequency effects vary within a grade?**

Our first aim was to describe the effect of lexicality, length, and frequency as a function of reading level in students with the same amount of education. Considering our anticipation of slower performance in the lexical decision (LD) task among participants with poorer reading proficiency, we will examine whether slowness alone accounts for the larger observed effects on response times (RTs) among less fluent individuals. It is worth noting that slower participants often exhibit more pronounced RT effects compared to faster participants. (Faust, Balota, Spieler, & Ferraro, 1999). If those effects occur at the decision stage, which is thought to involve a stochastic accumulation of evidence, then one would predict effect size to be proportional to the standard deviation of RTs across trials (Sigman & Dehaene, 2005). We will investigate this hypothesis using z-scores of the response time.

### **How is response to pseudowords affected by reading level?**

Another goal of our work was to better understand how students with different levels of reading ability process pseudowords. To this aim, we designed different types of word-derived pseudowords, also called “traps” because they must be rejected, in spite of their often close similarity to words. Much prior research has demonstrated an effect of the orthographic similarity of pseudowords to words (Davis & Bowers, 2006; Ferrand & Grainger, 1993; Grainger & Segui, 1990). This similarity is classically measured using Coltheart's N (Coltheart et al., 1977), which considers as “neighbors” two strings of the same length that differ by only one letter. Here, however, we explored a greater range of pseudoword types, based on the presence of transposed or mirror letters, as well as misspelled words.

Both developing and skilled readers can make errors characterized by internal letter transpositions (i.e., reading “from” as “form”) (Paterson et al., 2015). In LD tasks, children show a higher tendency in misclassifying pseudowords with transposed letters. This effect initially intensifies with reading



acquisition and subsequently diminishes to its minimum level among skilled readers. (Grainger et al., 2012). We expected our participants to be slower and less accurate when processing letter-transposition pseudowords than their double-substitution controls, an effect that would be the highest for our most fluent readers. Regarding mirror generalization, it is an early predisposition of the pre-reader's brain that must be inhibited or superseded when learning to read (Dehaene, 2009; Dehaene et al., 2015; Kolinsky et al., 2011; Pegado et al., 2011b, 2014). Indeed, identifying that two mirror letters are different is harder than differentiating between two non-mirror letters (Ahr et al., 2016). In addition, the sounds /p/-/b/ and /b/-/d/ are very close phonologically. Thus, we expected that processing pseudowords containing mirror substitutions, such as 'dateau' instead of the French word 'bateau', would require a greater effort than processing pseudowords arising from an equivalent, non-mirror letter substitution (e.g., 'fateau'; English equivalents would be 'dalance' [derived from 'balance'] versus 'falance'). As students with reading deficits often confuse letter-sound rules (Rack et al., 1992), we also introduced misspelled words that would sound like a word if the wrong GPC rule was applied (pseudohomophones). Prior research using lexical decision showed that students exhibited major difficulties distinguishing between words and pseudohomophones (Bergmann & Wimmer, 2008) which decreased in the course of reading development (Grainger et al., 2012). Thus, we expected higher RTs and error rates for orthographic traps than for control word approximations, and a reduction of this effect as fluency increased. Errors on pseudohomophones may be an outcome of poor reading experience exasperated by difficulties in learning French, a language with an opaque orthography. We hypothesized that all readers would be slower to classify pseudowords based on their distance to a real word, but that poor readers would be further penalized by pseudohomophones.

## **METHODS**

### **Participants**

Within the framework of the French national evaluations taken by all 6th-grade students, which marks

the commencement of middle school, we implemented our LD task in a student panel. These evaluations on expected grade-level French and Math abilities are done individually on computer, with the exception of a one-minute oral reading fluency test. The administration and data collection processes were overseen by the Direction of Evaluation, Prospective, and Performance (DEPP), the ministerial service responsible for education statistics in France (website: [www.education.gouv.fr/direction-de-l-evaluation-de-la-prospective-et-de-la-performance-depp-12389](http://www.education.gouv.fr/direction-de-l-evaluation-de-la-prospective-et-de-la-performance-depp-12389)).

The panel taking the LD task consisted of 3,472 students from schools chosen by the DEPP as representative of French population. The DEPP was also able to provide us with the reading fluency scores for 2,194 students. To examine the impact of reading ability on lexical decision results, we only kept students who had completed both tests. No differences in gender or socio-economic status between our initial and reduced sample were found. Out of these 2,194 students, only 1,501 completed the LD task entirely. We found that the lower the student's scores in fluency, the more they abandoned the lexical decision along the way. Their partial data were still collected, but the number of trials performed was too low to be analysed. Thus, the results described in this article concern the 1,501 students (806 girls and 695 boys, mean age = 11.0 y.o.) who had a complete LD dataset (120 trials) and reading fluency scores. This sample was still representative of the French population in terms of socio-economic status.

## **Reading fluency test**

Within the National Evaluations, student fluency was assessed by the standardized text, *Le Géant égoïste* included in the BALE battery (Jacquier-Roux et al., 2010). This text consists of 206 words, spread over 15 lines. The teacher administered this portion of the national evaluation individually. In a quiet room, the student was instructed to read the text aloud for one minute, as accurately as possible in normal reading speed. Teachers reported the student's number of words correctly read in one minute.

## **Lexical Decision Task**

The LD task was included in the computerized portion of the national assessment. Students worked individually with headphones in a group setting in the school's computer lab. The task started with written and oral instructions: "For this exercise, decide as fast as possible if what is written on the screen is a real word or a trap. Press M for a word and Q for a trap." This was followed by a video demonstration for each item category (word or trap). Students clicked a button when they were ready to start. Each item remained in the middle of the screen until the student responded by pressing 'Q' or 'M' or was skipped after a 5000ms time limit. Audio-visual feedback was provided. Positive audio feedback increased in tone with consecutive correct responses to encourage pursuit of winnings streaks. We collected measures of accuracy and response time (RT in ms). There were twelve different lexical decision modules, each composed of 120 stimuli, 60 words and 60 pseudowords (see below). Students were randomly assigned to a module. The whole task was administered in one block, and stimulus order was randomized within each student.

### ***Word stimuli***

We first extracted all mono-lemmatic and mono-morphemic words from the Lexique 3.83 database (New et al., 2004) with a length of four to eight letters and a frequency higher than three per million. We manually excluded all potentially offending, inappropriate or foreign words, thus resulting in a stimulus set of 3,656 words. These items were then separated into four different frequency bands: very frequent, frequent, rare, and very rare (see table 1 for details). 12 modules were designed using this database. For each module, we randomly selected 3 words from each frequency category and each length, resulting in a factorial design with length (5 levels, 4-8 letters) and frequency (4 levels) as factors, for a total of 60 words per module.

Frequency category	Length				
	4 letters	5 letters	6 letters	7 letters	8 letters
Very frequent (greater than 100 per million)	beau (nice)	avion (plane)	visage (face)	message (message)	déranger (disturb)
Frequent (40 to 100 per million)	acte (act)	usine (factory)	camion (truck)	étudier (study)	vaisseau (vessel)
Rare (10 to 40 per million)	vélo (bike)	alibi (alibi)	carnet (booklet)	complot (conspiracy)	éprouver (experience)
Very rare (3 to 10 per million)	cerf (deer)	maçon (bricklayer)	abolir (abolish)	stocker (store)	élégance (elegance)

**Table 1.** Characteristics of words stimuli and examples.

### ***Pseudoword stimuli***

Using only words from the “very frequent” category, we built six categories of pseudowords “traps”, for a total pool of 1,196 items. For each module, we randomly selected two pseudowords for each type and each length, thus resulting in a factorial design with pseudoword category (6 levels, see below) and length (5 levels, 4-8 letters) as factors, and 60 pseudowords per module. We describe each of the pseudoword categories below. Examples are presented in Table 2.

*Orthographic traps.* These were misspelled words that could, by an erroneous grapheme-phoneme correspondence, sound like real word. These pseudowords were manually built by selecting all words in our pool with a given rule-based grapheme-phoneme correspondence, then over-regularizing it. We focused on the letters s, c, and g, whose pronunciation varies with context in French. The letter ‘s’ most frequently sounds as /s/ (e.g., in “sale”) except when a single ‘s’ is sandwiched between vowels (e.g., in “base”), causing it to sound /z/. Similarly, letters ‘c’ and ‘g’ respectively sound as /k/ and /g/ when followed by the vowels ‘a’, ‘o’ and ‘u’, but as /s/ and /Z/ when followed by the letters ‘e’ and ‘i’. As a result of those rules, the pseudoword ‘ausi’ should be read /ozi/, and the

pseudoword 'bage' should be read /baZ/. A reader for whom these rules have not been consolidated might read these pseudowords as the words /osi/ (also) and /bag/ (ring)

*Word approximations.* Those were pseudowords entirely made of frequent French trigrams. To build them, we calculated the frequencies of all legal letter trigrams in our word pool. Only trigrams with frequency greater than 1/10,000 were retained. Pseudowords of 4 to 8 letters were then built solely from those frequent trigrams. We implemented a Markov process that (1) draws an initial trigram at random, in proportion to its frequency; (2) uses the last two letters to continue with the next trigram, again drawing randomly based on frequency, and so on. For example, the 4-letter pseudoword "arie" could be built using the frequent trigrams 'ari' and 'rie'. Actual words were excluded by software and human inspection.

*Letter transposition traps.* Pseudowords of this category were constructed by inverting two adjacent internal letters of a word. The transposed bigram was composed of either two vowels or two consonants. Only pseudoword items whose bigrams exceeded a frequency of 1/10,000 were kept. Bigram frequency was calculated using the same method described above for trigrams.

*Two-letter substitution traps.* This category, a control for transpositions, was built by substituting the same bigram in each transposed pseudoword with another random bigram with frequency higher than 1/10,000. Consonants were replaced by consonants, and vowels by vowels. All the resulting bigrams from this substitution were controlled to ensure that their frequency exceeds a threshold of 1/10,000.

*Mirror letter traps (mirroring of letters b d p q).* This category was generated by mirroring mirror letters in the following way:  $p \rightarrow q$ ;  $q \rightarrow p$ ;  $b \rightarrow d$ ;  $d \rightarrow b$ . Items were only kept if the transformation yielded a pseudoword.

*Single-letter substitution traps (substitution of letters b d p q).* In this category, a control for

the mirror substitution traps, letters b d p q were substituted with a non-mirror letter: p→g ; q→j ; b→f ; d→t.

Pseudoword traps	Length				
	4 letters	5 letters	6 letters	7 letters	8 letters
Orthographic traps	bage (bague = ring)	plase (place = place)	inciet (inquiet = worried)	ésaiyer (essayer = try)	difisile (difficile = hard)
Word approximations	atio	ouvoi	jamure	répoure	voicider
Transpositions	ceil (ciel = sky)	juene (jeune = young)	geurre (guerre = war)	pafrois (parfois = sometimes)	regadrer (regarder = look)
Double substitutions	cuol (ciel = sky)	jaine (jeune = young)	gairre (guerre = war)	pansois (parfois = sometimes)	reganger (regarder = look)
Mirror substitutions	aibe (aide = help)	qièce (pièce = room)	dateau (bateau = boat)	musipue (musique = music)	qrochain (prochain = next)
Single substitutions	aite (aide = help)	gièce (pièce = room)	fateau (bateau = boat)	musijue (musique = music)	grochain (prochain = next)

**Table 2.** Characteristics of pseudowords stimuli and examples.

## Data Analyses

We used the fluency score of the student as a measure of reading automatization. Trials with a response time (RT) exceeding 200ms were included in a linear mixed-effect model, incorporating fluency level (quintiles 1-5), lexicality (word, pseudoword), length (4-8 letters), and word frequency (4 levels ranging from very frequent to very rare) as fixed effects. Subject and stimulus were treated as random effects. We also ran our analyses but with fluency as a continuous variable. Similar results were obtained. As the variation in effects between the different groups, however, was clearly non-linear, all further analyses were conducted with the fluency quintiles as a 5-level factorial variable, rather than as a continuous variable.

We used the 'mixed' function from R's afex package with the following formula:

$$dv \sim X_1 * X_2 * \dots * X_n + (1|subject) + (1|stimulus)$$

$X_i$  represents the combination of our interacting fixed effects.  $dv$  is either accuracy of the answer or RT. For RT, we computed a classical mixed effect model (Baayen et al., 2008) while for accuracy we used a logistic mixed effect regression with the binomial link function (Jaeger, 2008). Significance was computed using the Satterthwaite's degrees of freedom for RT and the Likelihood ratio test (LRT) for accuracy. All analyses used a significance threshold of  $\alpha=0.05$ .

Follow-up analyses were done using the simple effect analysis where we split the data into subsets according to the modulating variable(s) and recomputed the model with only the remaining variable(s).

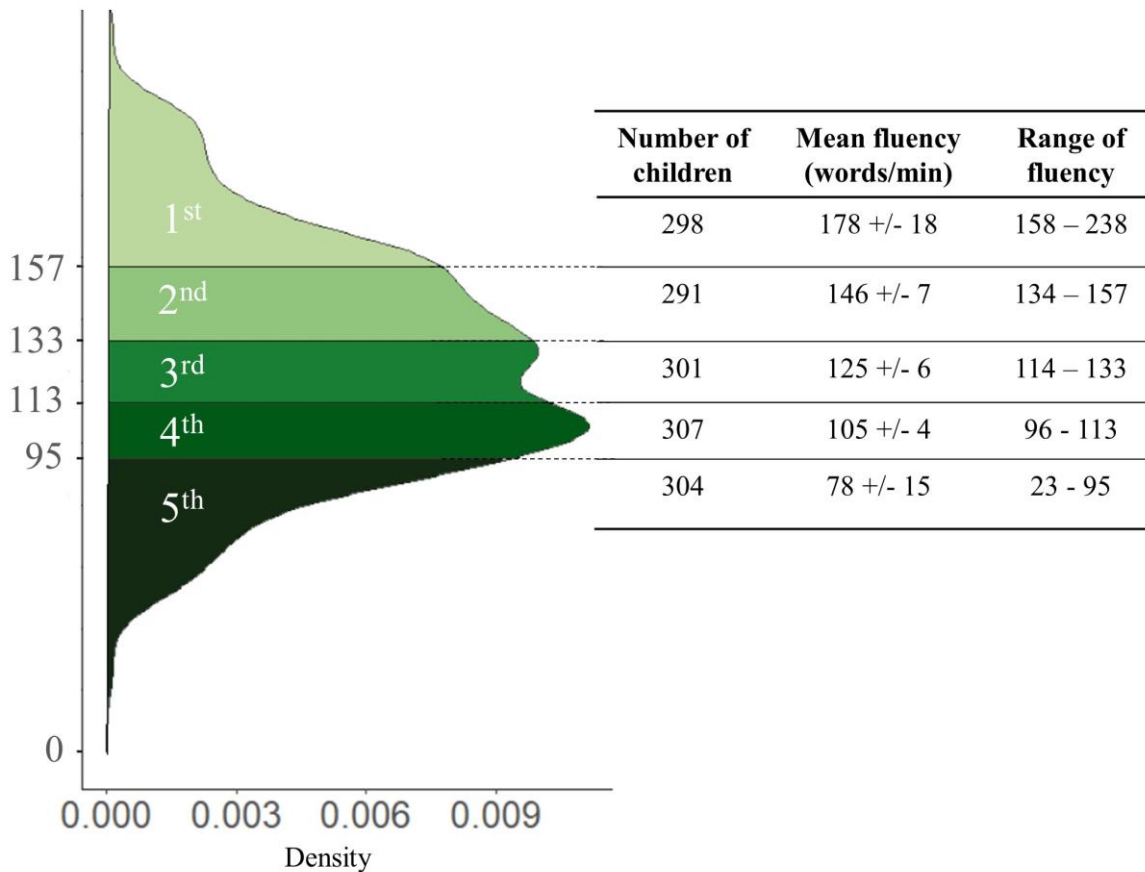
To check if the slowness of less fluent readers alone could explain why they showed larger (absolute) effects on RTs, we turned our raw RT into z-scores by subtracting from each subject's RT their overall mean and then dividing by their overall standard deviation (Zoccolotti et al., 2008), after verifying that our data fitted the conditions to apply Faust et al.'s (1999) rate and amount model (RAM).

## **RESULTS**

Figure 1 shows the distribution of text reading fluency in our sample, in words/minutes. Since the primary goal of our project was to assess within-grade variability in the size of the lexicality, length and frequency effects in the LD task results, we first separated students into five groups based on their fluency score, each quintile group containing 20% of the students. Fluency quintile 1 refers to the best readers and 5 to the poorest readers.

## Reading fluency

Number of words correctly read  
in one minute



**Figure 1.** Distribution of participant’s reading fluency on the National Evaluation. Reading fluency was assessed by measuring the number of words correctly read in a text in one minute. Color gradation from dark to light corresponds to fluency group levels, with light green corresponding to the most fluent students. Each quintile group represents 20% of the tested population.

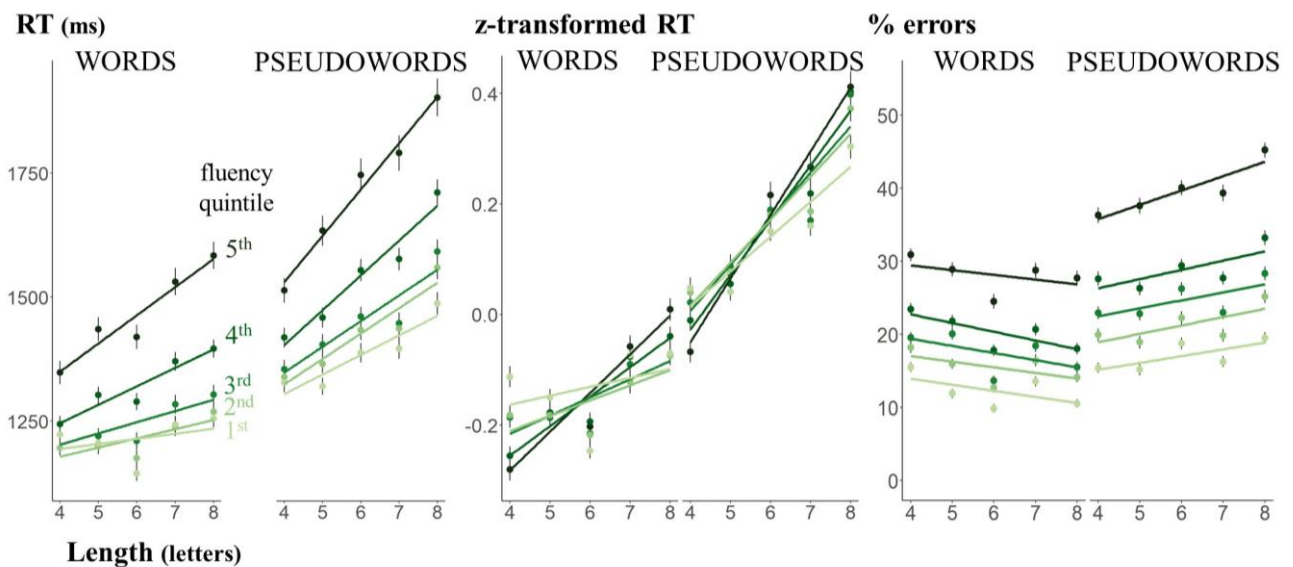
These quintiles are indicative of the expected Gaussian reading levels of pupils for this grade level according to the test's standardized norms. At the start of 6<sup>th</sup> grade students are expected to be capable of correctly reading 127.16 words per minute in this text, with a standard deviation of 29.41. Within our sample, 47.7% of the students perform above this norm, while 57.8% fall within the range of plus or minus one standard deviation. Moreover, 5.06% of the pupils exhibit reading abilities below two standard deviations.



## Length and lexicality effects

LD results are plotted in Figure 2 as a function of lexical status and length. In agreement with the literature on LD RT, responses to words were faster than to pseudowords,  $F(1,1063.1)=213.20$ ,  $p<.0001$ ; better readers read faster,  $F(4,1478.9)=58.86$ ,  $p<.0001$ ; RT increased with word length  $F(1,1062.5)=73.85$ ,  $p<.0001$ . All three two-way interactions were significant: fluency x lexicality,  $F(4,135933.4)=16.55$ ,  $p<.0001$ ; fluency x length,  $F(4,135826.6)=50.37$ ,  $p<.0001$ ; lexicality x length,  $F(1,1062.6)=7.63$ ,  $p=0.006$ . There was, however, no significant three-way interaction between fluency, lexicality, and length  $F(4,135833.3)=0.23$ ,  $p=0.92$ .

Performing simple effect analysis on the significant interaction between length and fluency, we found that that all quintiles exhibit a significant length effect. The steepness of the slope of this effect increased, as fluency level decreased. Looking at the interaction between length and lexicality, we found that both words and pseudowords exhibit a significant length effect, which slope was steeper for pseudowords than for words.



**Figure 2.** Length and lexicality effects on response times on correct answers (RT), Z-transformed RTs, and error rates. Each point represents the mean RT or error rate as a function of word length and fluency. Error bars represent one standard error of the mean. The slopes are the linear regression associated with the point.

As expected, less fluent readers were also slower in the lexical decision task. Could this slowness alone explain why less fluent subjects showed a larger (absolute) effect of word length on RTs? To answer this question, we turned our raw RT into z-scores. The results remain unchanged. Our mixed effect model on z-transformed RTs showed that all effects were significant except the three-way interaction,  $F(4,137202.4)=0.45$ ,  $p=0.77$ . Crucially, the fluency by length interaction remained significant,  $F(4,137202.4)=30.46$ ,  $p<.0001$ , indicating that the reduction of the length effect as fluency increased went beyond what could be due solely to faster responses. Fluent readers are not just faster, but have a genuinely smaller length effect.

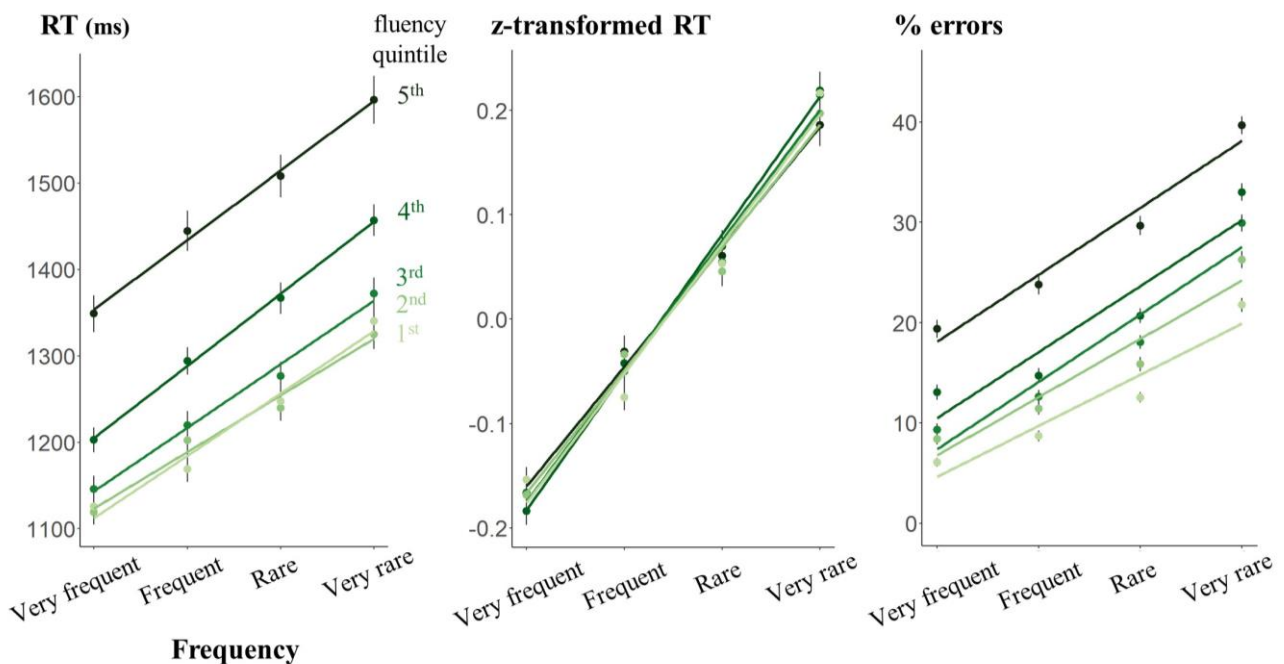
We next turned our attention to comparing error rates. Results mirrored those for RT. All subjects were more accurate for words than pseudowords,  $\chi^2(1)=91.85$ ,  $p<.001$ , and better readers were more accurate than their counterparts,  $\chi^2(4)=595.89$ ,  $p<.001$ . The main effect of length was not significant,  $\chi^2(1)=0.01$ ,  $p=0.92$ , but a significant interaction was found between length and lexicality,  $\chi^2(1)=14.16$ ,  $p<.001$ : greater length increased the likelihood of errors for pseudowords, but decreased it for words. We also found a significant interaction between fluency and lexicality,  $\chi^2(4)=21.02$ ,  $p<.001$  though not between fluency and length,  $\chi^2(4)=9.36$ ,  $p=0.053$ , nor a three-way interaction,  $\chi^2(4)=1.00$ ,  $p=0.91$ . In the absence of an interaction between fluency and length, we conclude that across all trials, none of the quintiles exhibits any significant length effect.

## **Frequency effect on words**

Subsequently, we focused our investigation on word items within the LD task, specifically examining the influence of word frequency and its relationship with fluency variations. Correct RTs and error rates sorted by the four bands of frequency appear in Figure 3. Both RT and accuracy were affected by a significant main effect of frequency (RT:  $F(1,561.66)=101.40$ ,  $p<.0001$ ; accuracy:  $\chi^2(1)=219.05$ ,  $p<.001$ ). Again, there was a main effect of fluency quintile (RT:  $F(4,1461.01)=54.23$ ,  $p<.0001$ ; accuracy:

$\chi^2(4)=422.22$ ,  $p<.001$ ) and a significant interaction between fluency and frequency (RT:  $F(4,70582.79)=4.59$ ,  $p=0.001$ ; accuracy:  $\chi^2(4)=30.93$ ,  $p<.001$ ) due to the fact that the slope of the frequency effect decreased in the most fluent readers. Analyses restricted to each fluency quintile showed that, for both RT and accuracy, performance for all fluency quintiles dropped with less frequent words. The slope of decreased performance was only less steep for our best performers.

Again, we used z-transformed RTs to examine whether the reduction in the frequency effect as fluency increased was solely due to faster overall responses. The main effect on frequency remained significant,  $F(1,620.63)=116.36$ ,  $p<.0001$ , but the two-way interaction vanished,  $F(4,71977.60)=1.06$ ,  $p=0.37$ . This finding suggests that, once the speed of their responses was considered, the frequency effect on RT was actually identical for all students, no matter their fluency level. In other words, frequency appears to be a variable that affects the decision stage and whose amplitude therefore covaries with the standard deviation of RTs, while word length does not.



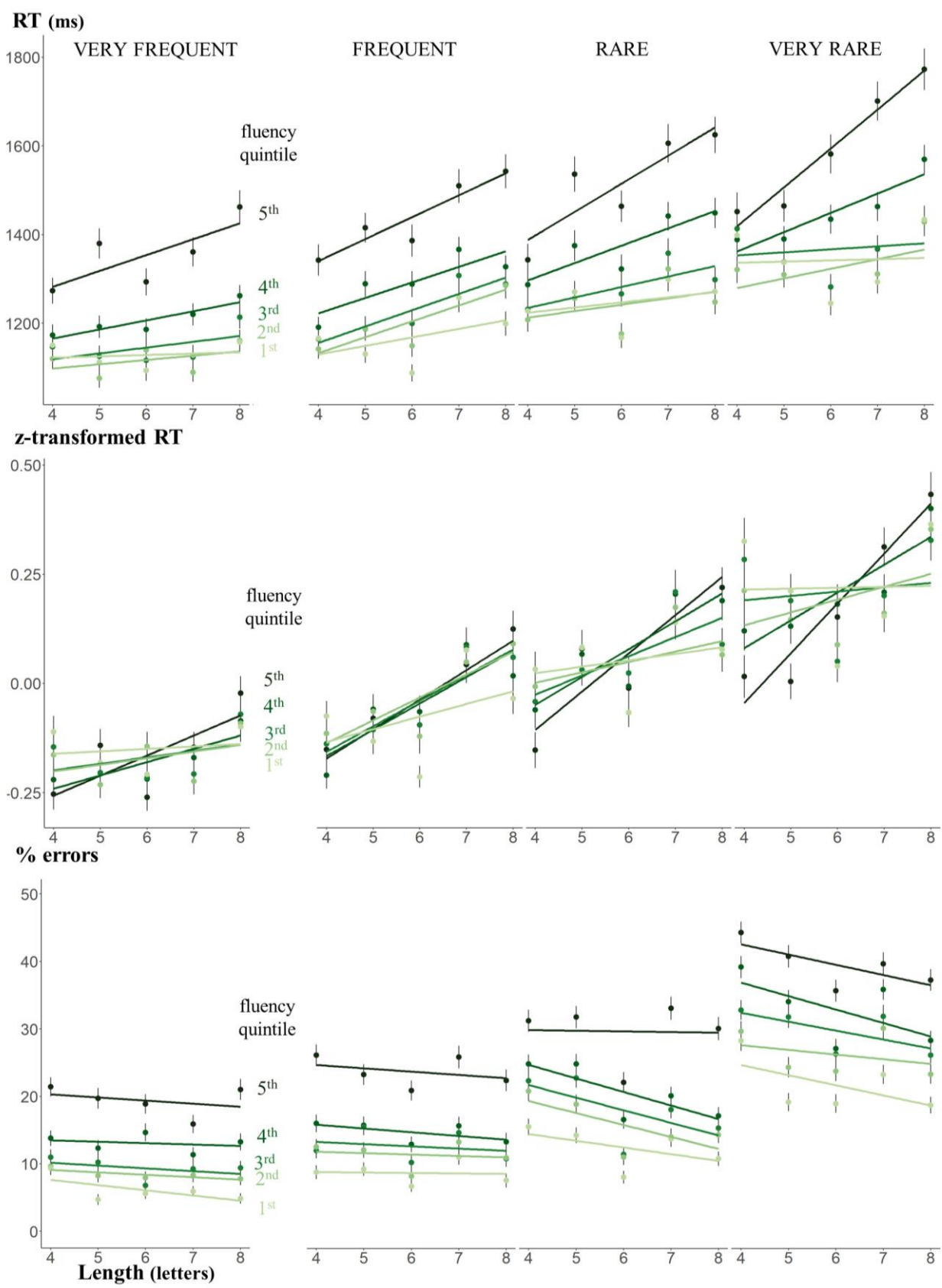
**Figure 3.** Frequency effects on response times on correct answers (RT), Z-transformed RTs and error rates. Each point represents the mean RT or error rate as a function of frequency and fluency. Error bars represent the standard error of the mean. The slopes are the linear regression associated with the points.

## Interaction between length and frequency

Our large sample also allowed us to investigate the interaction between length, frequency, and fluency (figure 4). Our prediction was that these variables should have a 3-way interaction on RTs because (1) fluent readers would show little or no length effect, regardless of frequency, as they consistently rely on the fast lexical route; (2) less fluent readers would show an increasingly marked length effect as word frequency decreases, because lower frequency decreases the probability that they use the lexical route.

In a general linear model restricted to words only, the predicted three-way interaction between fluency quintile, word frequency and length was significant for RT, though not for accuracy (RT:  $F(4,70564.14)=5.59$ ,  $p=0.0002$  / accuracy:  $\chi^2(4)=2.76$ ,  $p=0.60$ ). Figure 4 shows that RTs were affected by the frequency x length interaction reported earlier, and that, as predicted, this effect decreased as fluency increased. A simple effect analysis on length, separately for each fluency x frequency level, showed a clear trend: the most fluent students exhibited no length effect, no matter the frequency of the words, whereas the least fluent student showed a significant length effect even for very frequent words.

Results on z-transformed RTs were almost exactly the same as for RT, except that, as expected, there was no longer a significant main effect of fluency quintile,  $F(4,71970.80)=1.30$ ,  $p=0.27$ , as well as any significant interaction between word frequency and fluency,  $F(4,71979.29)=0.99$ ,  $p=0.41$ , as previously reported. The interaction between length and frequency was also not significant,  $F(1,598.16)=0.95$ ,  $p=0.33$ . As the three-way interaction between length, word frequency and fluency level remained significant, our conclusions remained unchanged.



**Figure 4.** Interaction between length, frequency and fluency on response times on correct answers (RT), Z-transformed RTs and error rates. Each point represents the mean RT or error rate as a function of length,

frequency and fluency. Error bars represent the standard error of the mean. The slopes are the linear regression associated with the points.

## Processing of pseudowords

We designed our pseudoword traps for two-by-two comparisons: orthographic traps versus words approximations, transpositions versus double substitutions, and mirror versus single substitutions.

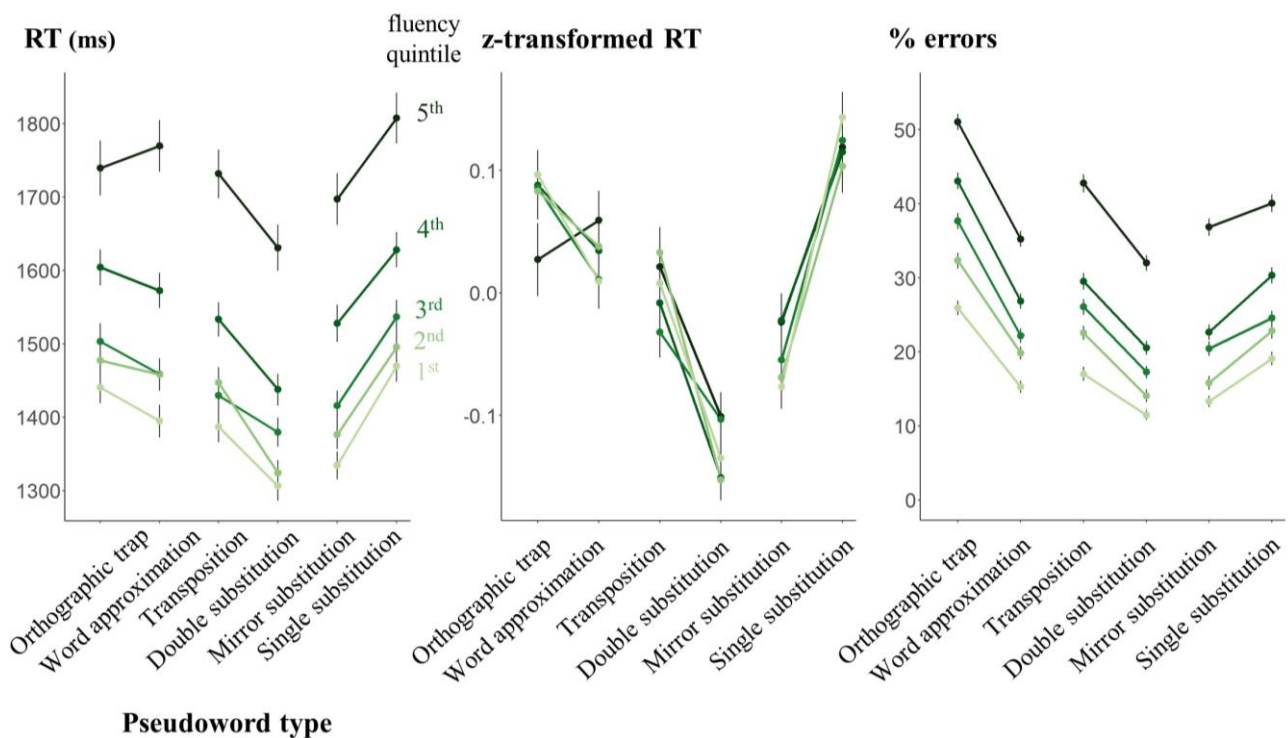


Figure 5 shows the data for these comparisons, which we consider in turn.

**Figure 5.** Responses to pseudowords: reponse time on correct answers (RT), Z-transformed RTs and error rates.

### *Impact of orthographic traps*

Our mixed effect analysis only showed a significant main effect of pseudoword type on accuracy,  $\chi^2(1)=32.35$ ,  $p<.001$ , but not on RT,  $F(1,147.36)=0.038$ ,  $p=0.85$ , meaning that when they were correct, students were equally fast to classify orthographic traps and word approximations. We also found a significant effect of fluency quintile on both RT and accuracy (RT:  $F(4,1428.30)=37.10$ ,  $p<.0001$  ; accuracy:  $\chi^2(4)=393.42$ ,  $p<.001$ ) and a significant interaction between fluency and pseudoword type

only for RT (RT:  $F(4, 19168.23)=3.10$ ,  $p=0.015$  ; accuracy:  $\chi^2(4)=2.80$ ,  $p=0.59$ ), which was preserved in z-transformed RTs,  $F(4,20407.0)=2.81$ ,  $p=0.024$ . Performing simple effect analysis, we found that this interaction highlighted that word approximations tended to be classified faster than orthographic traps, except in the least fluent students.

### ***Effect of letter transpositions***

Results of our mixed effect models confirmed our expectations. For both RT and accuracy, we found a main effect of pseudoword type (RT:  $F(1,175.21)=11.60$ ,  $p=0.0008$  ; accuracy:  $\chi^2(1)=26.70$ ,  $p<.001$ ) meaning that transpositions were harder to classify than double substitutions. Initially, we found that the better the reader, the better the overall performance as highlighted by the main effect of fluency, (RT:  $F(4,1403.38)=46.75$ ,  $p<.0001$ ; accuracy:  $\chi^2(4)=403.63$ ,  $p<.001$ ), in the absence of an interaction with pseudoword type (RT:  $F(4,21294.44)=2.02$ ,  $p=0.088$  ; accuracy:  $\chi^2(4)=3.53$ ,  $p=0.473$ ). However, looking at z-scores, we did find a significant interaction between type and fluency level,  $F(4,22646.1)=2.46$ ,  $p=0.043$ , confirming that the difference observed between the two type of pseudowords is steeper as fluency level increases.

### ***Processing mirror substitutions***

Surprisingly, our results on this type of pseudoword departed from our prediction that mirror substitutions should be more difficult than single substitutions (figure 5). Our mixed effect model showed a significant main effect of pseudoword type on both RT and accuracy (RT:  $F(1,174.64)=11.04$ ,  $p=0.0011$  ; accuracy:  $\chi^2(1)=13.73$ ,  $p<.001$ ), but with a difference in favor of mirror substitutions. There was a significant main effect of fluency quintile (RT:  $F(4,1422.50)=46.80$ ,  $p<.0001$  ; accuracy:  $\chi^2(4)=383.65$ ,  $p<.001$ ) and a significant interaction on accuracy only (RT:  $F(4,20941.65)=0.20$ ,  $p=0.94$  ; accuracy:  $\chi^2(4)=22.02$ ,  $p<.001$ ). Within each fluency quintile, faster and more accurate performance with mirror letters reached significance for all but the least fluent students,  $\chi^2(1)=3.23$ ,  $p=0.072$  -- the converse of our predictions.

What could be the underlying reason of this surprising effect? One possibility is that mirror confusions are more likely to access the mental lexicon, such as the activation of 'balance' when encountering 'dalance'. Subsequently, top-down feedback from the lexicon might facilitate the detection of the erroneous letter 'd', not to be confused with the 'b'. Control pseudowords like 'falance' would not benefit from this letter correction to lexical input. To probe the contribution of the lexical route, we tested for an effect of the frequency of the original word. There was no main effect of frequency (RT:  $F(1,173.64)=0.20$ ,  $p=0.65$  ; accuracy:  $\chi^2(1)=0.12$ ,  $p=0.72$ ) nor, crucially, its interaction with pseudoword type (RT:  $F(1,172.55)=0.08$ ,  $p=0.78$  ; accuracy:  $\chi^2(1)=0.67$ ,  $p=0.41$ ). Thus, the lexical route does not appear to contribute much to the processing of those pseudowords, if at all.

A second possibility is that as some mirror-letter substitutions violated the orthographic statistics of French, such violations could have facilitated the rejection of those pseudowords. Indeed, in French, the letter 'q' is very rare and is almost always followed by the letter 'u'. However, when p was substituted with a q, this graphotactic rule was violated in all but one of our mirror-substituted pseudowords containing a 'q'. To test this idea, we removed all items with a 'q' substitution and repeated our mixed effect analysis. The main effect of pseudoword type vanished for both RT and accuracy (RT:  $F(1,133.21)=2.19$ ,  $p=0.14$  ; accuracy:  $\chi^2(1)=1.43$ ,  $p=0.231$ ). There was also no interaction with fluency on RT,  $F(4,15227.15)=1.84$ ,  $p=0.12$ , and a minor one on accuracy:  $\chi^2(4)=12.87$ ,  $p=0.012$ . This finding suggests that our paradoxical effect on RT (more efficient processing of mirror substitutions) was in fact entirely due to the specifics of the letter 'q' – and mirror letters did not pose specific difficulties for our participants.

## **DISCUSSION**

Our primary goal was to evaluate the lexical decision performance of 6<sup>th</sup> grade students at different levels of reading fluency, a known predictor of school reading comprehension (Fuchs et al., 2001a; Hudson et al., 2005a; National Reading Panel, 2000; Pinnell et al., 1995). Our results for reading ability



within a single grade converge with prior research comparing LD performance across grades, children with adults, or normal readers with dyslexics – but the originality of the present work is to thoroughly characterized the variability in a large sample of students within a single grade year (6<sup>th</sup> grade). Our findings corroborate teachers’ reports of a large diversity of reading ability when students arrive in their first year of middle school. Based on student fluency scores, the best performing quintile had already reached adult reading levels, reading more than 200 words per minute, while participants in the worst quintile lagged behind 3<sup>rd</sup> grade expectations, a striking result with practical implications (Andreu et al., 2022). National statistics suggest that, depending on the definition of literacy, more than 10% of the French population is struggling with reading, and more than 5% remains functionally illiterate when they leave school and enter active life (Chabanon & Rosenwald, 2020). The present results suggest that the LD test in 6<sup>th</sup> grade may pick such difficulties at a moment where they might still be acted upon.

Our findings are largely compatible with the hypothesis of two different pathways for reading words and pseudowords (Castles, 2006; Coltheart et al., 2001b; Di Filippo et al., 2006b). In line with previous research, we found a main effect of lexicality, with faster responses to known lexical items. The presence of a significant length effect on pseudoword RTs fits with the hypothesis that pseudowords are deciphered via a slow sublexical route. The decrease in the length effect as student's fluency level increases provides an estimate of the effectiveness of their procedure: the better the students are at reading, the faster their lexical and sublexical pathways.

The impact of length on pseudoword reading contrasts with what was reported by Juphrad and colleagues who found no length effect on pseudoword RT in skilled readers (Juphrad et al., 2004a). The type of pseudowords used in the tasks can possibly explain the difference. Juphrad et al's pseudowords were closest to our trigrams, most of which are orthographically distant from real words, while their words were of very high frequency (minimum 134 per million). Thus, it was far easier for their readers to base their decisions on lexicality, which may explain the absence of a length effect on pseudowords’

RT. In our study, however, the presence of rare words, as well as the frequent proximity of pseudowords to real words, may have forced participants to rely more heavily on decoding reading procedures.

Access to the mental lexicon was evaluated by varying the frequency of the word stimuli. As expected, we found a significant frequency effect on both RT and accuracy. The absence of an interaction with fluency in the z-score analysis indicated that, once corrected for greater slowness and variability of responses in the less fluent students, the frequency effect had the same size in all fluency quintiles. This result is in line with those observed by Burani et al (Burani et al., 2002) in younger children, where they showed a frequency effect on naming latencies but no significant interaction between word frequency and grade.

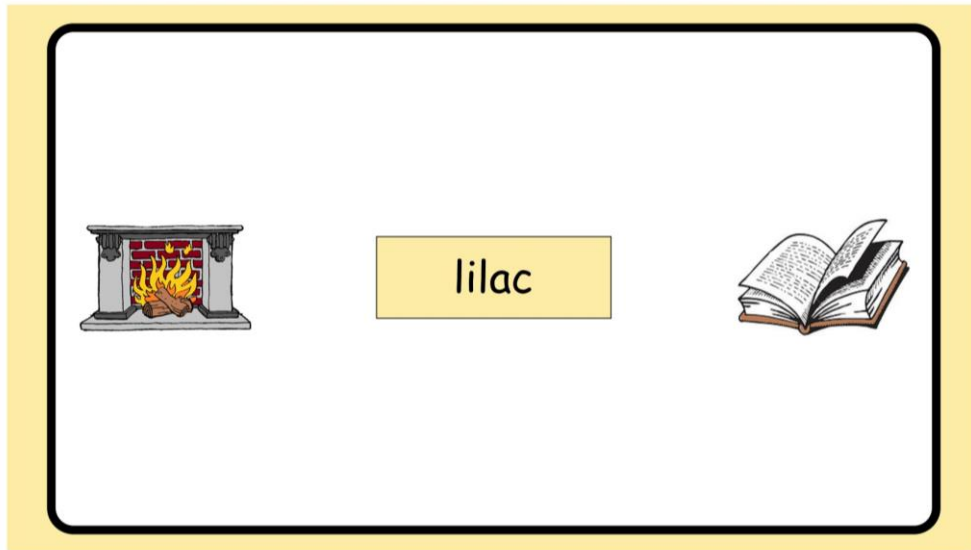
Delving deeper, we found RT interactions between length and frequency to be dependent on fluency. All readers except the least fluent ones used lexical strategies to read very frequent words. The findings for our best readers regarding lexicality, length and frequency effects were consistent with those described by Araujo et al in Portuguese 3<sup>rd</sup> to 5<sup>th</sup> graders (Araújo et al., 2014b). Our poorest reader, however, were similar to previously reported cases of dyslexic readers, as evident by a length effect on RT, thus betraying a strategy of accurate but sublexical reading even for frequent words (Araújo et al., 2014b; Zoccolotti et al., 2005b). In other words, our poorest readers managed to correctly judge a majority of frequent words but did so by first identifying them through a slow sublexical reading process.

The second goal of our study was to examine the impact of orthographic distance on pseudoword judgments. In general agreement with previous research, we found that stimuli that were most similar to real French words yielded that highest error rate (Bergmann & Wimmer, 2008; Grainger et al., 2012). Going further, we systematically compared orthographic traps with trigram-based approximations, transpositions with double substitutions, and mirror with single substitutions. We found greater mistakes on orthographic traps (homophones) than on trigram controls. Such errors are mainly due to

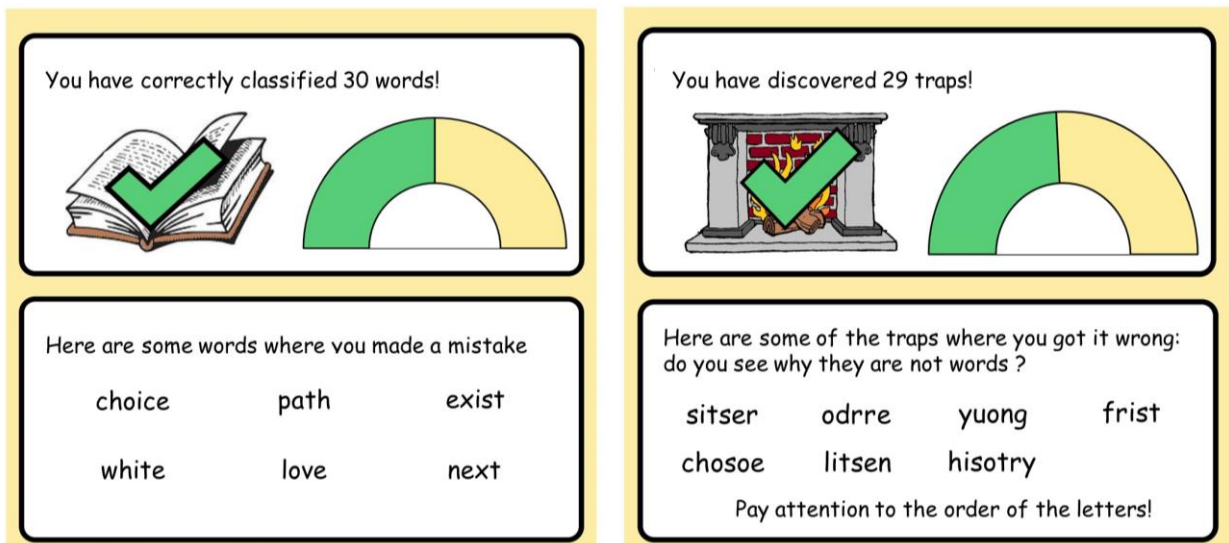
regularizations of letter sounds, which is a marker of the use of the sublexical procedure. In line with prior research, our results revealed that transposed letters led to increased response times (RT) and error rates when compared to letter substitutions (Chambers, 1979; Grainger et al., 2012). Notably, this effect became more pronounced as we compared the least fluent readers to the most fluent ones. This observation can be interpreted as an indication of the growing reliance on the lexical procedure, where all letters are processed simultaneously. As a result, there is a higher likelihood of letter position confusions occurring. Finally, looking at mirror and single substitutions, we found a surprising effect: mirror substitutions were faster and more accurately classified than single substitutions. The effect appears to be mainly due to the transformation  $p \rightarrow q$ . When removing these pseudoword items, the difference between these two traps also disappears. We conclude that mirror letters, relative to substitutions, do not pose difficulties for French middle-school readers.

From a practical point of view, these results support the use of LD as a complementary test to oral reading fluency, one that can reveal details of the reading processes in students who show obvious difficulties on the fluency test (Balota et al., 2006; Seidenberg & McClelland, 1989). For example, our poorest readers showed a length effect on RTs to frequent words, suggesting that they know these words but identify them using a sublexical procedure, similarly to dyslexic students (Araújo et al., 2014b; Castles, 2006). This insight from LD, which could not be obtained from fluency alone, may help flag students requiring special intervention. To this aim, we have begun to introduce teachers and students with a gamified version of our LD test (figure 6). At the end of this version of the test, students and teachers are provided with a summary of the main results, broken down into two categories: first, the results on words, with examples of errors made in the most frequent category; and second, the results on pseudowords, with examples of the student's main error type. A tip on how to avoid these types of errors in the future is added at the bottom of the page. In future research, we plan to test the usefulness of this gamified lexical decision test for both students and teachers.

## Test screen



## Personalized feedback



**Figure 6.** Gamified version of the lexical decision task. Students are asked to send each stimulus into the dictionary if it is a word, and into the fireplace if it is a trap. At the end of the session, they receive feedback on their performance, including the number of errors on words and pseudowords and examples of errors. For words, errors are sorted by their frequency. For pseudowords, only the stimuli from the category in which the student made the most errors are displayed. A tip on how to avoid these traps is given.

There are several limitations to this work. As the Ministry of Education carried out the recruitment, we were not able to formally exclude dyslexic students from our participants. Thus, it is likely that part of the group of students with the greatest difficulties was in fact composed of dyslexics. Similarly, we

were not able to check for the presence of other learning disabilities that might have disrupted the test, such as attention deficit disorder. We were also not able to test students twice or more, to assess test-retest reliability, neither on the LD task nor on the fluency test. Furthermore, the great diversity of items presented to each child left little room for more than 2 or 3 repetitions of the same condition (2 repetitions of each length-type combination for pseudowords, 3 repetitions of each length-frequency combination for words). The high error rates on pseudowords therefore made it impossible to analyze the length effect within each type of pseudoword.

## **CONCLUSION**

In summary, our study showed that lexical decision results variation within a grade in middle school are the same than those observed in longitudinal studies. We suggest that LD cannot replace other oral reading tests such as fluency, but provides an additional assessment that sheds detailed light on the efficiency of the two reading routes, i.e. lexical and sublexical reading procedures. LD has now proven its correlation with other tasks of oral reading. Given its ease of use, LD could prove an excellent tool to differentiate between readers in need of extra practice, and those with more serious deficits marked by an inability to establish an efficient mental lexicon.

# **CHAPTER 2 – THE MARIETTE, A SCREENING TEST FOR READING ERRORS IN PRIMARY SCHOOLS**

## **ABSTRACT**

In the realm of dyslexia research, the umbrella term "dyslexia" encompasses a spectrum of reading difficulties characterized by diverse types of errors exhibited by readers. Unfortunately, prevailing reading assessment tools in France predominantly focus on evaluating reading speed and accuracy while neglecting the critical aspect of quantifying error types. To address this significant gap, we present the Mariette test, a novel reading aloud assessment featuring a 294-word nonsense text. The Mariette test incorporates regular, irregular, and pseudo-words, along with strategically embedded word traps. These elements enable the identification of various dyslexic profiles that may manifest during the orthographic-visual analysis stage as well as in lexical and sublexical reading. In this study, we standardized the outcomes of reading time, accuracy, and error types by administering the Mariette test to 812 French pupils spanning from 1st to 5th grade, encompassing all five years of elementary education in France. Our research elucidates how these outcomes evolve throughout primary school in typically developing readers. The Mariette test, coupled with associated normative standards, facilitates the establishment of thresholds indicative of reading fragility and deficit. To enhance the utility of the Mariette test, we provide comprehensive guidelines for interpreting the diverse types of errors observed. Additionally, we offer recommendations for supplementary assessments aimed at refining the diagnosis of dyslexia. We posit that the Mariette test has the potential to provide educators and practitioners with a more comprehensive understanding of dyslexic profiles within the elementary school context, thereby contributing to more effective support and intervention strategies for students with dyslexia.

## **INTRODUCTION**

Developmental dyslexia, a prevalent learning disability, impacts a significant portion of the population, with prevalence rates ranging from 3% to 12%, contingent upon language and specific diagnostic criteria (Di Folco et al., 2022; Lindgren et al., 1985). Identifying this condition at the earliest juncture and subsequently providing tailored support, especially to young children, represents a paramount endeavor to maximize students' academic success in school.

In the realm of French language assessment, there exists a variety of test batteries designed to identify written language disorders. These assessments encompass both longstanding tools like Alouette (1967), and more contemporary offerings such as Evaléo (2018). These batteries, developed by a collaborative effort between researchers and practitioners, encompass a range of tests aimed at evaluating different facets of written language proficiency. Most of these assessments are underpinned by the Dual Route Model of reading (Coltheart, 2007), which posits that reading involves a dynamic interplay between two distinct pathways for processing written information. The first pathway, referred to as the lexical pathway, is responsible for recognizing words that are already stored in the reader's mental lexicon. Within this lexicon, words are organized based on their frequency, resulting in more common words being processed more swiftly than their less frequent counterparts. Essentially, this pathway facilitates the rapid reading of familiar words that the reader has encountered previously. Conversely, the second pathway, known as the sublexical pathway, is instrumental in decoding new words or pseudowords by applying grapheme-phoneme conversion rules. This pathway operates at a slower pace, as evidenced by the robust length effect, wherein longer words necessitate more time for accurate reading compared to words stored in the mental lexicon. Remarkably, for words spanning 2 to 8 letters, there is a consistent reading response time observed for those stored in the mental lexicon, demonstrating that word length does not significantly impact reading time (Coltheart, 2007). Before these two pathways come into play, readers engage in a visual-orthographic analysis stage. During this initial phase, readers identify the constituent letters



composing the word and establish associations between these letters, their respective positions within the word, and their interactions with neighboring words, a process particularly pertinent in text reading (Friedmann & Coltheart, 2018).

To test reading, most batteries contain exercises in reading aloud and identifying written words. The exercises designed to measure the automation of reading involve reading texts over different durations (1 minute: BALE, 3 minutes: Exalang 8-11, Alouette-R, 5 minutes: Exalang 11-15, Vol du PC). In some batteries, these texts are significant (BALE, Evaléo, Exalang 8-11, Exalang 11-15, Vol du PC), while in others they are not (Alouette-R, Evaléo). Some of these texts also include pseudowords to assess decoding skills (Exalang 11-15, Vol du PC). All rely on measures of completion time and number of errors to establish scores, to which are sometimes added measures of pseudoword error rate where applicable.

Written word identification exercises are designed to assess more precisely the functionality of the reading pathways. Composed of lists of words varying in frequency, length, and regularity, as well as pseudowords, they measure (using generally over a hundred items) reading time and the number of errors for each type of exercise presented (BALE, BELEC, Exalang 5-8, Exalang 8-11, Exalang 11-15, LMC-R). The results obtained enable a differentiation to be made between surface dyslexia (impairment of the lexical pathway, marked by many errors in the reading of irregular words) and phonological dyslexia (impairment of the sublexical pathway, marked by many errors in the reading of pseudowords).

Combined with the results of reading comprehension exercises, these tests claim to allow practitioners to decide on the functionality of reading and on the presence of surface dyslexia, phonological dyslexia or mixed dyslexia (meaning both reading pathways are impaired). However, this categorization has its limits, as there are many other cases of dyslexia reported in the literature. Below we briefly describe several other types of dyslexia found in the research, that are not taken into account by traditional screeners in French:

- **Letter identity dyslexia**, refers to a deficit in encoding the abstract identity of letters. This is a rare form of dyslexia, in which patients are unable to read because they are unable to identify letters correctly. They cannot name a letter, identify a written letter by its name or sound, or match upper and lower case letters (Brunsdon et al., 2006). A deficit in the output of the visuo-orthographic analysis affects all the functions of this analyser and leads to visual dyslexia, a profile in which the patient omits, substitutes or adds letters in words, whatever the letter and whatever its position (Friedmann et al., 2012).<sup>2</sup>

- **Attentional dyslexia**, a deficit in binding letters to the words they appear in, resulting in migrations between words (for example, reading 'win fed' as 'fin fed', Shallice & Warrington, 1977). In this type of dyslexia, letters are correctly identified and almost always retain their relative position, but they migrate to a neighboring word - for example 'dig fin' is read 'fig fin'. Migration can occur both horizontally and vertically, most often at the end of a word, and in French more frequently from the second word to the first (indicating that it is not simply a matter of perseveration of a letter or phoneme). They can take the form of substitutions, additions or omissions of letters present in neighboring words. Attentional dyslexia has been broadly described in a case of an English reader (Rayner et al., 1989) and its properties examined in detail in multiple cases in Hebrew (Friedmann, Kerbel, et al., 2010a; Lukov et al., 2015), Arabic (Friedmann & Haddad-Hanna, 2014a), and in French (Potier Watkins et. al., 2023);

- **Letter-position Dyslexia**, refers to readers with a selective impairment in the ability to encode the relative position of letters in words. Letter migrations occur most often between two inner, adjacent letters, particularly when the error also forms a word - for example, 'from' is read as 'form', or 'smile' as 'slime'. These errors occur both in reading aloud and silently, in any task that involves identifying or

---

<sup>2</sup> Not published yet, our team has recently identified a child with letter identity dyslexia in France.

understanding a word. Cases of selective developmental letter-position dyslexia have been described in readers of Hebrew (Friedmann, Dotan, et al., 2010; Friedmann et al., 2015; Friedmann & Rahamim, 2007, 2014); English (Kezilas et al., 2014; Kohnen et al., 2012); Arabic (Friedmann & Haddad-Hanna, 2014b, 2012), Turkish (Güven & Friedmann, 2019), Italian (Traficante et al., 2021), and French (Potier Watkins et al., 2023);

- ***Neglect dyslexia (or “neglexia”)***, the reader may omit and substitute letters in reading, but errors predominantly affect one side of the word (i.e., reading ‘yellow’ as ‘pillow’, an example from Ellis et al., 1987). This impairment may affect reading exclusively, without neglect of other visual stimuli. Neglexia has been documented in developmental cases with children (Friedmann & Nachman-Katz, 2004; Nachman-Katz & Friedmann, 2007, 2010).

- ***Orthographic-visual analyzer dyslexia***, (also called Visual dyslexia in the literature) is characterized by errors where the incorrect response shares some letters with the target word. For example, reading 'unicorn' as 'united,' 'coronary,' or 'acorn.' These errors may involve substitutions, omissions, additions, or transpositions of letters. Notably, these errors are not attributed to issues with letter positioning, attentional dyslexia, or individual letter identification (in contrast to letter identification dyslexia, where single letter recognition remains intact). Furthermore, these errors cannot be explained by neglect dyslexia. It appears that this type of dyslexia arises from a deficiency in either the output of the orthographic-visual analyzer or the orthographic input buffer. (Friedmann et al., 2012; Friedmann & Coltheart, 2018).

- ***Vowel Dyslexia***, characterized by omissions, migrations, substitutions, and additions of vowel letters, with no corresponding errors in consonant letter reading. This condition primarily affects the sublexical route. In cases of pure vowel dyslexia, vowel errors are limited to pseudowords. When individuals also exhibit surface dyslexia alongside vowel dyslexia, they are compelled to read real words through the sublexical route, potentially resulting in vowel letter errors in real words. Notably, individuals with

vowel dyslexia do not typically encounter difficulties in manipulating vowel phonemes during oral language tasks (Khentov-Kraus & Friedmann, 2018; Güven & Friedmann, 2021).

- **Deep dyslexia** is a complex condition characterized by relying exclusively on the semantic lexicon when reading. In this case, the reading process involves recognizing a word, comprehending its meaning, mentally visualizing it, and then naming what they've understood or imagined. This multi-step approach leads to semantic errors, which are a hallmark of deep dyslexia (e.g., reading 'sand' as 'beach'), as well as morphological and visual errors. Given the readers reliance on meaning-based reading, individuals with developmental deep dyslexia face heightened challenges when encountering abstract words, morphologically complex words, and function words. Additionally, they exhibit a profound difficulty in reading pseudowords. It's important to note that developmental deep dyslexia is relatively rare and has primarily been documented in the context of English language research. (in children with Williams Syndrome, Temple, 2003, 2006), in Japanese (Yamada, 1995), and in Arabic (Friedmann & Haddad-Hanna, 2014b).

- **Grapheme-phoneme conversion dyslexia**, often confused by most screeners in France as phonological dyslexia, is a deficit in reading tasks that depend upon grapheme-phoneme decoding such as reading novel words and pseudowords, with even the conversion of single syllables impaired. Readers who cannot learn through the tradition grapheme-phoneme conversion route are required to memorize words by their orthographic form (Castles & Coltheart, 1993). This dyslexia was first discussed in the case of a participant with dyslexia who was able to read familiar words fluently, but unable to read pseudowords despite of an intact ability to orally repeat and write spoken items (Beauvois & Derouesne, 1979).

- **Phonological output buffer dyslexia** involves a deficiency in the short-term memory buffer responsible for holding phonological units until they are ready for production and for assembling phonemes into words. Individuals with aphasia stemming from brain damage and afflicted by this deficit encounter challenges when reading aloud lengthy and morphologically intricate pseudowords

(e.g., 'binking') or real words (e.g., 'dancing'). Since the phonological output buffer plays a role not only in reading aloud but also in speech production, those with a deficiency at this stage also exhibit comparable difficulties in oral tasks such as repeating complex stimuli. Additionally, they tend to substitute numbers and function words with other words from the same category. (Dotan & Friedmann, 2015; Guggenheim & Friedmann, 2014).

The list provided above, while not exhaustive, serves as an adequate introduction to the objective of this study, which aims to address various forms of dyslexia. It is worth noting that these deficits can occur in isolation, implying the independence of each stage. However, they can also co-occur within the same individual.

In her research, the dyslexia specialist Dr. Friedmann advocates for the development of new diagnostic tools. The concept involves creating a reading screener that goes beyond the conventional lists of irregular words and pseudowords, by including carefully selected "trapping" stimuli designed to elicit each type of error particular to different dyslexia types. For instance, in some individuals, only the presentation of "transposable" words like "from-> form" or "smile->slime" can reveal errors in letter positioning that might otherwise go unnoticed. Consequently, because dyslexia might only manifest as a reading slowdown or even an absence of deficits on standard tests. (Kohnen et al., 2012; Potier Watkins et al., 2023). It is therefore essential that the standards for these new tests consider not only the number of errors in each category, but also the precise type of error made by the child, by precisely characterizing each of his errors.

With this in mind, our laboratory, in partnership with the CERENE schools and diagnostic center, has developed a word identification test, the Malabi, standardized for 6<sup>th</sup> and 7<sup>th</sup> graders (Potier Watkins et al., 2023), and a test of text reading aloud, the Mariette, normed for pupils from 1<sup>st</sup> to 5<sup>th</sup> grade, to which this chapter is devoted. Based on observations of multiple types of dyslexia, the Mariette test has been standardized in terms of speed, accuracy and also in terms of the different types of errors made by the child, in order to characterize any difficulties as accurately as possible. The aim is to

provide a rapid tool that, in just a few minutes of reading aloud, gives the practitioner an initial overview of the main reading difficulties that a child may face. A test such as this also enables the child's reading to be assessed under the most ecologically friendly conditions possible, because unlike a list of words, reading a text requires good use of the oculomotor mechanisms. Using this test alone, however, is not enough to conclude what type of dyslexia the child might have. It provides some initial clues, but should be followed up with other, more in-depth tests.

This chapter explains the design of the Mariette test and presents the results obtained on almost 800 normal-reading pupils. This work is still in progress, and a final section will present our on-going questions to be taken up in post-doc work after this thesis.

## **METHOD**

### **Design of the Mariette**

The Mariette is a nonsense text of 294 words to be read aloud. The words were divided into 4 paragraphs, matched in terms of mean word length (Kruskal-Wallis test,  $p=0.45$ ) and frequency (Kruskal-Wallis test,  $p=0.62$ ). Following the progression of difficulty of grapheme-phoneme correspondences in French, described by Potier Watkins et al. (Potier Watkins et al., 2019), the text progressively introduces increasingly difficult graphemes. For example, the complex digram "gn" appears for the first time in the third paragraph, whereas the easier digram "ou" is introduced as early as the first paragraph. To examine the influence of text alignment (left-aligned or justified) on pupils' reading speed, the first two paragraphs were left-aligned, while paragraphs 2 and 4 were justified.

The text includes many words or pseudowords which, if they are subject to a specific type of error, lead to the production of another French word. There are pairs of words such as 'cape page' or 'dont vent' which are likely to give rise to attentional errors, because the exchange of letters leads to other words ('cage pape'; 'dent vont'); words or pseudowords whose letters can be transposed, such as 'fiable', 'arbi' or 'rogné', i.e. where the transposition of two letters leads to another word;

pseudowords such as 'truche', which are likely to give rise to lexicalisations; irregular words such as 'parfum' or 'femme' for regularisation errors; and words containing contextual rule traps such as 'dragée' or 'cousin', which are likely to be read as 'draguée' or 'coussin' in the event of poor mastery of French contextual rules. Each paragraph contains five irregular words and four pseudowords, and the text contains at least ten examples of each of the other types of error.

The fact that this test is nonsense is essential in the detection of dyslexia, as it makes it impossible for the pupil to rely on the context to guess the words that follows and avoid the pitfalls, which they can do on meaningful texts (Friedmann & Rahamim, 2007a). The full text is available in the Appendix.

## **Participants**

We have standardised this test for primary school pupils, because the aim was to identify pupils' reading difficulties as early as possible in the course of schooling. To do this, 812 pupils from 1<sup>st</sup> to 5<sup>th</sup> grade (age: 5 years 11 months to 11 years 11 months) were recruited in 2023 from 7 schools of diverse backgrounds, spread across 4 different academies: 38% of the pupils tested lived in rural areas and 27% of the pupils tested came from priority education, in order to make our sample more representative. However, these percentages were slightly higher than at national level, where 20% of pupils are educated in priority education (Stefanou, 2022) and 20% of pupils are living in rural areas (Direction de l'Evaluation, de la Prospective et de la Performance, 2021).

These pupils were all tested in March 2023 and the 1<sup>st</sup> and 2<sup>nd</sup> graders were also tested twice more, in December 2022 and June 2023, each time over a period of 2 to 4 weeks, in order to take account of the significant progression in their reading ability over the year. The exact number of pupils tested for each test session are shown in Table 1.

1.8% of pupils were excluded because their age was greater than or less than 2 standard deviations from the mean, and 2.4% of the remaining pupils were excluded because their speed or success rate was more than 3 standard deviations from the mean, resulting in a sub-sample of 778 pupils from 1<sup>st</sup> to 5<sup>th</sup> grade (age: 6 years 5 months to 10 years 8 months).

<b>Grade</b>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>
<b>December session</b>	84	79	/	/	/
<b>March session</b>	80	86	73	100	103
<b>June session</b>	88	85	/	/	/

**Table 1.** Number of pupils per grade for each measurement period

## **Procedure**

Parents were informed of the purpose of our study and of the possibility for them to object to their child's participation via a form given to the pupil by the teacher. The tests were carried out by our research team during school time, in a quiet room close to the classroom. During the test, the child was asked to read aloud the Mariette text, printed on an A4 paper, Calibri 14 font, 1.5 line spacing, for a maximum of 5 minutes.

The pupil was informed that the text made no sense and contained pseudowords, with the following instructions: "This text is a story that makes no sense, so it is normal if you don't understand it. In this story, there are words that don't exist, so it is normal if you don't know them. I'm going to ask you to read this story out loud, as best you can, and to go as far as you can in 5 minutes". Once the child had understood the instructions, the page was turned over in front of the pupil and the stopwatch started to record the reading time for each paragraph and the total reading time. In order to record errors on each word and guarantee the quality of the analyses of the pupils' verbal productions, each pupil's performance was recorded so that it could be listened to again.

During the test, pupils were allowed to follow the text with their finger. Pupils, especially the younger ones, received encouragement from us to keep them motivated by the exercise. So as not to discourage pupils at the start and middle year of 1<sup>st</sup> grade, only half the text was initially presented to them. If the pupil reached the end of the second paragraph in less than 5 minutes, the second half of the text was then offered. When marking, self-corrections and word skips were considered as errors. Occasionally, students skipped an entire line. When this happened, the number of words in the omitted line was subtracted from the total number of words read by the student.



To assess the validity of this test, 20% of randomly sampled pupils also took the Alouette test (Lefavrais, 2005), a test that has already been standardised and is widely used by paramedical professionals and researchers to screen for dyslexia. The Alouette text was presented in the same format as the Mariette, typed in Calibri 14 with 1.5 line spacing. The images in the original text were not reproduced. The Alouette text was always presented after the Mariette text.

## **Error characterisation**

### ***Percentage of errors on different types of words***

As mentioned in the introduction, in order to study different types of errors, most test batteries and articles in the literature refer to error rates in a given category of words (regular words, irregular words, pseudowords, pseudohomophones, transposable words, etc.). So, in order to compare our results with those in the literature, we first calculated for each pupil and each level the average success rate for regular words (16 words in the text containing neither digrams nor silent endings and not being functional words), irregular words (22 words in the text), pseudowords (16 words in the text), transposable words (20 words in the text) and word pairs (9 pairs in the text).

### ***Typology of each of the student's errors***

The previous classification does not, however, take account of the exact nature of the error made by the student, and does not include all the words in the text, even though all of them were the subject of at least one error. To remedy this, we then transcribed each of the pupil's errors phonetically. When the child made multiple responses to all or part of the stimulus, it was his first response to each syllable that was transcribed - for example, when the word '*fiable*' was read as '*fiad... ah no, fiable*', the pupil's response was transcribed as '*fiadle*' in order to capture as much of his initial production as possible. In order to characterise the type of error more precisely, we then applied a procedure similar to that described by Potier-Watkins et al. for the Malabi tests (2023). The different types of error we considered and some examples are presented in Table 2. Certain types of error present in the Malabi

have not been reported here due to their rarity. These errors have been included here under the label *'other'*.

For errors that could be attributed to letter additions, omissions or substitutions, the origin of these errors could then be refined, between attentional errors, resulting from migrations of letters between words, letter position errors, resulting from migrations of letters within the word or visual errors, which include all addition, omission or substitution errors that could not be attributed to one of the other two types. An addition, omission or substitution error could be categorised as attentional if the letter or letters involved maintained the same relative position as in the neighbouring words (first letter, last letter, inner letter). Here, we considered as neighbours the words surrounding the target word (for example, the neighbours of the word *'gradée'* are the words *'plante'*, *'dix'*, *'places'*, *'fiables'*, *'pose'*, *'mare'*, *'la'*, *'sur'* and *'étrange'*). An addition, omission or substitution error in which the letter in question was already present in the word could be categorised as a letter position error. Other additions, omissions or substitutions of letters could be attributed to visual errors. When an error made by a child was too far removed from the original word to be interpreted (the word *'sur'* read as *'sorti'*, for example), it was categorised under the label *'other'*.

This classification is not unequivocal. In fact, it was often the case that the same error could be attributed to several different types without it being possible to decide (for example, the word "nous" read as "sous" could be the result of a substitution within functional words, or of a substitution of the letter "n" by the letter "s", which itself could be either an attentional substitution or a letter position substitution). Other errors could be associated with a succession of different errors (for example, it was plausible that *'pose'*, read as *'possé'*, was a contextual rule error followed by a substitution). These different types of error were all taken into account in the standards, with the underlying idea that if a child suffers from a specific dyslexia, he or she should make many more errors in one category than in another, and that it is the accumulation of a certain type of error and the exceeding of the thresholds set by the standards that should alert the practitioner. For example, a child with letter position dyslexia

should have more errors of this type than the norm for his grade but should remain within the norms for attentional errors and other visual errors. A child with visual dyslexia should show a large number of visual errors but could also show a large number of attentional and letter position errors, because visual dyslexia is linked to a defect in the visuo-orthographic analysis, which also includes the association of a letter with its position in the word and between words. In all cases, it is the combined use of several increasingly precise tools that will enable the diagnosis to be refined. The Mariette could be followed by the Malabi to confirm these initial observations.

<b>Error type</b>	<b>Description</b>	<b>Stimuli sensitive to that type of error</b>	<b>Example → possible error</b>
<b>Attentional</b>	Addition, omission or substitution of letters present in neighbouring words. Letters that migrate keep their relative position within the word. These migrations can take place vertically or horizontally. The words surrounding the target word are considered to be neighbours.	All text items and in particular word pairs in which the migration of a letter between neighbouring words that retains its position within the word creates another existing word (cape-page / puis-sois).	<i>Addition</i> Page dira → page dirage <i>Omission</i> plante étrange → plate étrange <i>Substitution</i> lame rime → rame rime
<b>Letter position</b>	Letter transposition within words and pseudowords. Omission or substitution of an instance of a doubled letter, or doubling of a letter	All items in the text, and in particular migratable words and pseudowords: items for which a transposition in the word forms a word.	<i>Transposition</i> crique → cirque arbi → abri <i>Addition</i> sardine → sardrine <i>Omission</i> prirent → pirent <i>Substitution</i> depuis → dupuis
<b>Visual</b>	Omissions, substitutions and additions of letters resulting neither from a letter position error nor from an attentional error.	All text items	<i>Addition</i> mare → marche <i>Omission</i> compris → copris <i>Substitution</i> lange → lande
<b>Regularisation</b>	Regularisation of phonemes in irregular words	Irregular words	second → /səkō/

<b>Pronunciation of final silent letters</b>	Regularising the sound of the final silent letters	Words ending in a silent letter	nuit → /nuɪt/
<b>Mispronunciation of digrams</b>	Regulating the sound of digrams	Words containing digrams	plante → /plant/
<b>Incorrect application of contextual rules</b>	Confusion of contextual rules /g/ or /ʒ/ for the letter g - /s/ or /z/ for the letter s - /s/ or /k/ for the letter c - /s/ or /t/ for the letter t	Words containing a g, a c, a t an s or two s internally.	regarde → /rəʒard/
<b>Substitution within functional words</b>	Substitution of one functional word (determiners, pronouns, prepositions and conjunctions) by another	Functional words (determiners, pronouns, prepositions and conjunctions)	le → les, dans, du
<b>Difficulty with long and morphologically complex words</b>	Simplification of long words with complex morphology	Morphologically complex words	marcherions → marchons

**Table 2.** The different types of error recorded in Mariette

As not all the pupils read the same number of words during the test, the quantification of the different types of error was measured in terms of the percentage of errors made in relation to the number of words likely to result in such an error. For example, only irregular words can give rise to regularisation. A child who encountered 4 irregular words during the test and made 3 regularisations would have a regularisation percentage of 75%. The closer this score is to 100%, the more the child has made the same error each time he or she has encountered a certain type of word. The words counted for each type of error are described in Table 2.

## Other observables

### ***Fatigability***

Fatigability is a very useful parameter to estimate, as dyslexic students make more effort to read and therefore tire more quickly than normal readers. Here we characterised it as the variation in number of words correctly read per minute (called speed hereafter) from one paragraph to the next measured as a percentage using the following formula:

$$\text{difference paragraph 1} - \text{paragraph 2} = \left( \frac{\text{speed paragraph 2}}{\text{speed paragraph 1}} - 1 \right) \times 100$$

A negative difference indicates a decrease, while a positive difference indicates an increase. Only differences for which at least 80% of pupils had reached the end of the paragraphs involved were reported.

### ***Self-corrections***

The term self-correction refers here to the fact that a student makes a mistake and then corrects himself. In some tests, such as the Alouette, self-corrections are considered as correct answers, with the underlying idea that the student's hesitations will necessarily result in a slowdown in their reading speed.

As we are looking here at the exact errors made by the pupil, it is vital to keep a record of their first verbal production, even if they subsequently self-correct. This is why, in the Mariette, we consider any incorrect answer to be an error, and why we have taken self-corrections into account at a later stage. As not all the pupils read the same number of words, we expressed these self-corrections as the proportion of self-corrected errors in relation to the total number of errors made.

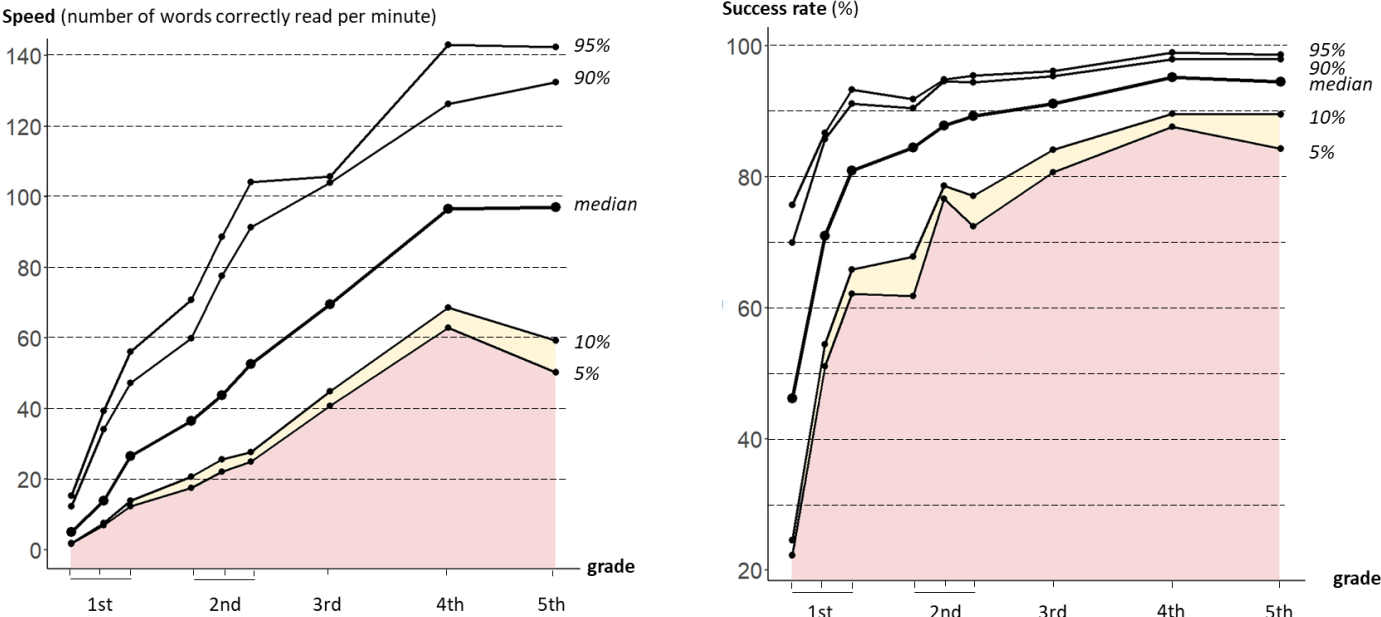
## **RESULTS**

### **Reading speed and success rate**

Figure 1 shows how pupils' reading speed and success rate changed over the course of schooling. In order to situate pupils among their peers, coloured thresholds are also shown in this figure: yellow for scores between the tenth and fifth percentiles, and red for scores below the fifth percentile. Between the start of 1<sup>st</sup> grade and the end of 4<sup>th</sup> grade, reading speed, expressed in words read correctly per minute, increases almost linearly ( $r=0.79$ ,  $p=0.021$ ), reaching a plateau between 4<sup>th</sup> and 5<sup>th</sup> grade. Pairwise comparisons showed that reading speed increased significantly between each level ( $p<0.05$  for each comparison) except between 4<sup>th</sup> and 5<sup>th</sup> grade ( $t(196.2)=-0.133$ ,  $p=0.89$ ).

The success rate followed a less linear profile: it rose sharply during the first year of primary school, from 46.6% to 79.5%, an increase underlined by the significance of pairwise comparisons ( $p<0.001$

between the beginning, middle and end of 1<sup>st</sup> grade). This increase then continued at a more moderate pace between the beginning of 2<sup>nd</sup> grade and 5<sup>th</sup> grade, reaching 93.9% success rate. Pairwise comparisons testify to this slowdown, since they are non-significant between the end of 1<sup>st</sup> grade and the beginning of 1<sup>st</sup> grade ( $t(163.6)=-1.69, p=0.093$ ), between the middle of 2<sup>nd</sup> grade and the end of 2<sup>nd</sup> grade ( $t(165.7)=0.035, p=0.97$ ), and between 4<sup>th</sup> and 5<sup>th</sup> grade ( $t(199.4)=1.35, p=0.18$ ).

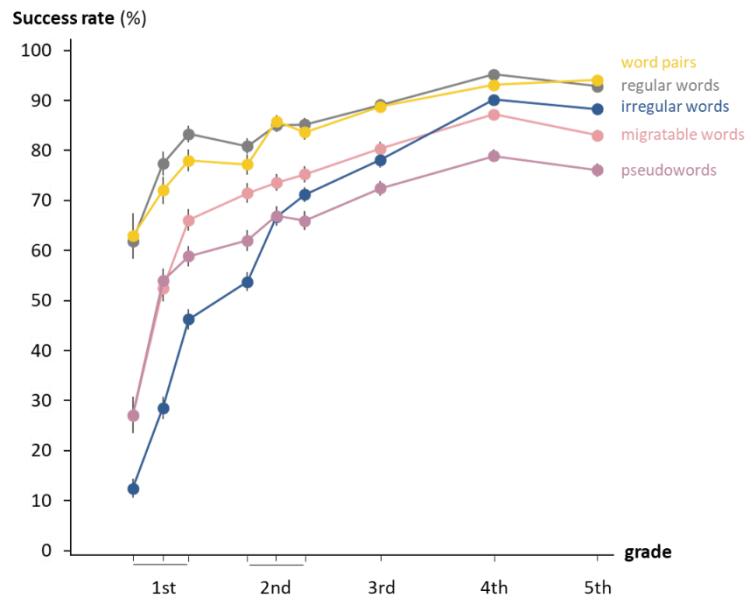


**Figure 1.** Changes in reading speed and success rate over the course of schooling. Colours are showing the thresholds at which students can be considered fragile (score below the 10th percentile) and impaired (score below the 5th percentile).

**Success rates for different word categories**

An initial measure of the types of errors made by the child consisted, for each grade, in averaging across pupils the rate of success in different word categories (regular words, irregular words, pseudowords, transposable words and word pairs). The results are shown in Figure 2. An analysis of variance on the percentages of success on regular words, irregular words and pseudowords showed a regularity effect (better success on regular words than on irregular words) as well as a lexicality effect (better success on regular words than on pseudowords), which were significant at all grades, even for the youngest pupils. A second analysis of variance showed that the difference between regular words and transposable words was also significant for all grades, with the success rate on regular words always higher than the success rate on transposable words. A final analysis of variance showed that

there was no significant difference between regular words and word pairs, except at the end of 1<sup>st</sup> grade ( $F(1,85)=255.30, p=0.021$ ), where the success rate for word pairs was slightly lower than the success rate for regular words.



**Figure 2.** Changes in the percentage of success for different types of stimuli (regular words, irregular words, pseudowords, transposable words and word pairs) over the course of schooling.

However, using only the success rate to compare results in these different word categories masks major disparities. For example, among children at the start of 1<sup>st</sup> grade, only 19.2% of errors on transposable words are transpositions, 37.7% of errors on word pairs were attentional errors and 38.1% of errors on irregular words were regularizations. While these percentages increase with grade, in 5<sup>th</sup> grade, transpositions account for only 67.9% of errors on transposable words, attentional errors account for 75% of errors on word pairs and regularizations account for 59.4% of errors on irregular words. It is therefore difficult to equate type of word with type of error, even for the oldest pupils. To get as close as possible to the pupils' difficulties, it therefore seemed important to us to also take an interest in pupils' exact errors, especially as the rate of errors characterized as 'other' was close to 0% from the middle of 1<sup>st</sup> grade, a sign that the method used here was able to characterize almost all the errors made by the pupils.

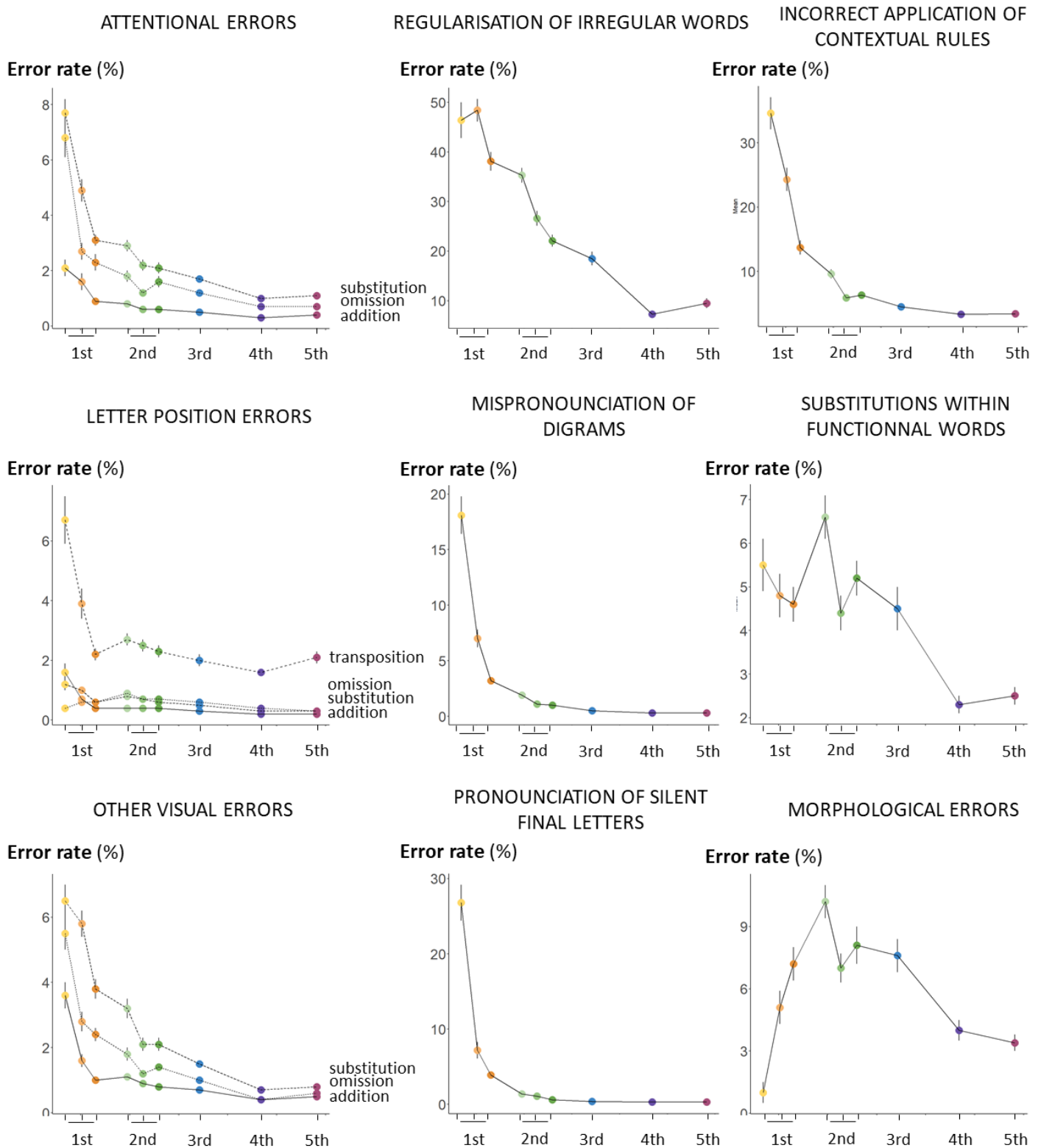
## Error rate for each error type

An analysis of student error rates for each type of error, shown in Figure 3, reveals a decrease in errors in each category between the start of 1<sup>st</sup> grade and 5<sup>th</sup> grade. However, this decline does not follow the same pattern depending on the type of error considered.

We observed a similar trend in the rates of attentional and visual errors: they fall sharply during the first year of primary school and then continue to fall more slowly from 2<sup>nd</sup> to 5<sup>th</sup> grade, where they are close to 0%. In other words, by the end of primary school, pupils make virtually no errors of this kind. The trend is slightly different for letter position errors, which also fall sharply between the beginning and end of 1<sup>st</sup> grade. While additions, omissions and substitutions of letters already present in the word disappear at the start of 2<sup>nd</sup> grade, transpositions reach a plateau at around 1.2% of errors. This is partly linked to the transposable pseudowords contained in the text (such as 'arbi' or 'oviles'), which lead to transpositions in the oldest pupils (in 1<sup>st</sup> graders 3.9% of transpositions were made on these pseudowords, a percentage that increases with schooling to reach 42.5% in 5<sup>th</sup> grade). While the mispronunciation of digrams and the pronunciation of silent endings fades very quickly over time, virtually disappearing from the beginning of 2<sup>nd</sup> grade, the incorrect application of contextual rules decreases over time but still peaks at around 4% from the middle of 2<sup>nd</sup> grade to 5<sup>th</sup> grade, a sign that these contextual rules are still not perfectly mastered at the end of primary school. Regularisation is very frequent at the start of 1<sup>st</sup> grade, where almost 40% of irregular words read have been regularised. The amount of regularisations increased between the beginning and middle of 1<sup>st</sup> grade and then decreased steadily from the end of 1<sup>st</sup> grade to 5<sup>th</sup> grade, even though the rate remains high: in 5<sup>th</sup> grade, an average of 10% of irregular words encountered are still regularised, probably due to the presence of infrequent irregular words such as 'dolmen'. Morphological simplifications follow a bell curve: errors increase between the beginning of 1<sup>st</sup> grade and the beginning of 2<sup>nd</sup> grade, then decrease from the beginning of 2<sup>nd</sup> grade to the middle of 5<sup>th</sup> grade. The increase in errors during the first year of primary school is probably an artefact, linked to the small number of words that can generate this type of error. Finally, confusion of functional words decreases during the first year of primary school



before increasing at the start of the second year of primary school and then decreasing again over the course of schooling, although the error rate remains fairly high, at around 5%.



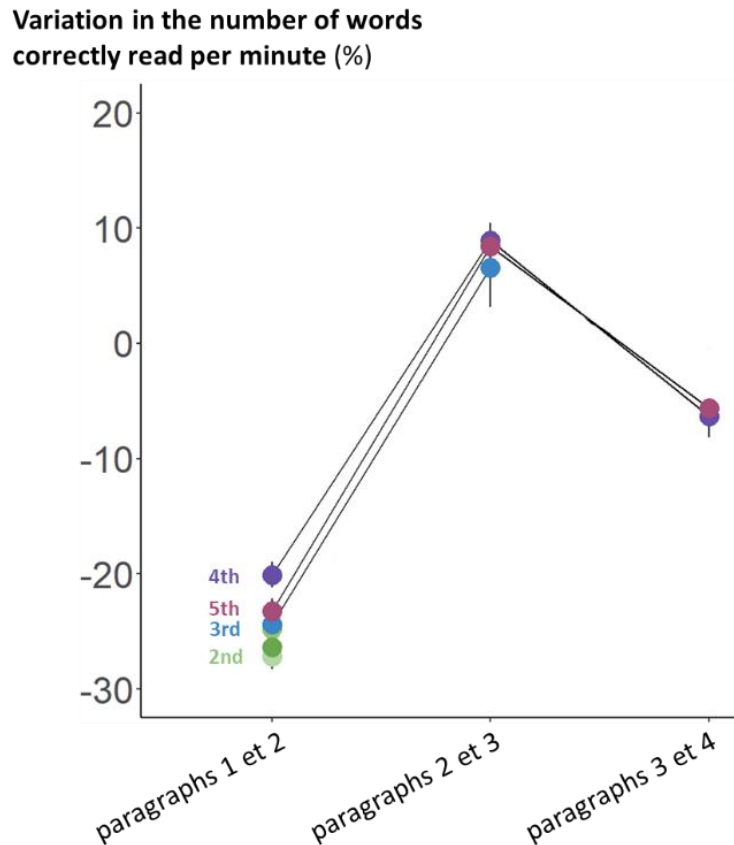
**Figure 3.** Evolution of different types of error over the course of schooling. These percentages were calculated in proportion to the number of possible occurrences of the error. For example, an error percentage of 100% on

regularisations means that every time the child read an irregular word, he made a regularisation. This is why the sum of the percentages in the graphs above does not equal 100.

## **Other observables**

### ***Fatigability***

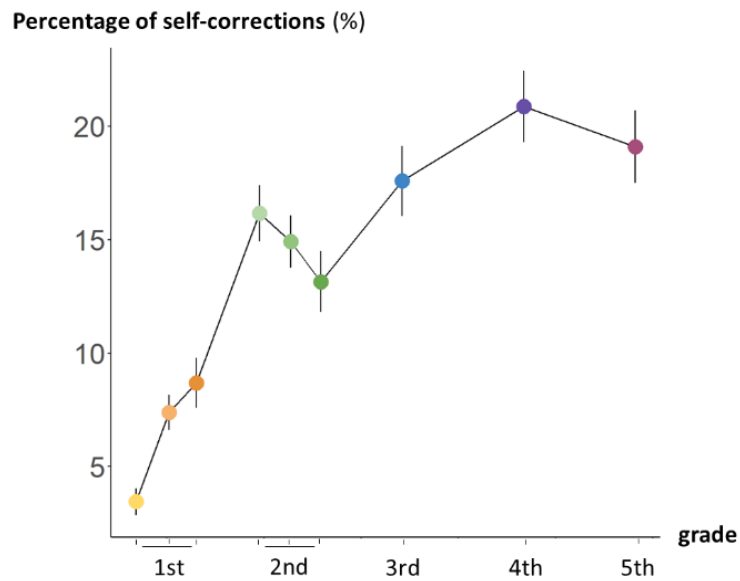
Fatigability, measured as the variation in speed between two paragraphs, is graphically described over time for each grade in Figure 4. The following variations can be observed: the pupils all slowed down between paragraph 1 and paragraph 2 (they read between 20% and 27% slower, depending on the grade considered) as well as between paragraph 3 and paragraph 4 (a drop in speed of around 6% for all the grades considered). On the other hand, between paragraph 2 and paragraph 3, the average speed increased for all the grades considered. Since the paragraphs are matched in terms of frequency and average word length, one possible interpretation of this result could be a facilitating effect of aligning the text on the left, and conversely, a deleterious effect of a justified presentation, as paragraphs 1 and 3 were left-aligned, whereas paragraphs 2 and 4 were justified. However, this interpretation should be taken with caution insofar as other parameters such as the frequency of bigrams or the number of neighbours were not measured for each of the Mariette words, and it is highly likely that the paragraphs differ on these points, which could also have had an influence on the variations observed.



**Figure 4.** Variations in reading speed (number of words correctly read per minute) between different paragraphs. These variations were only calculated for grades where at least 80% of students had reached the end of the paragraphs in question. A negative value indicates a slowdown between the two paragraphs, while a positive value indicates an acceleration.

### ***Self-corrections***

As shown in Figure 5, self-corrections are rare at the start of 1<sup>st</sup> grade, with only 3% of errors being self-corrected. This percentage increases over the course of 1<sup>st</sup> grade, and even more sharply between the end of 1<sup>st</sup> grade and the beginning of 2<sup>nd</sup> grade, when it rises from 8% to 16%. It then falls during 2<sup>nd</sup> grade, before rising again in 3<sup>rd</sup> grade and stabilising at around 20% between 4<sup>th</sup> and 5<sup>th</sup> grade. These 20% self-corrections are mainly due to traps that pupils fall into, such as pseudo-transposable words, which they then correct instantly.

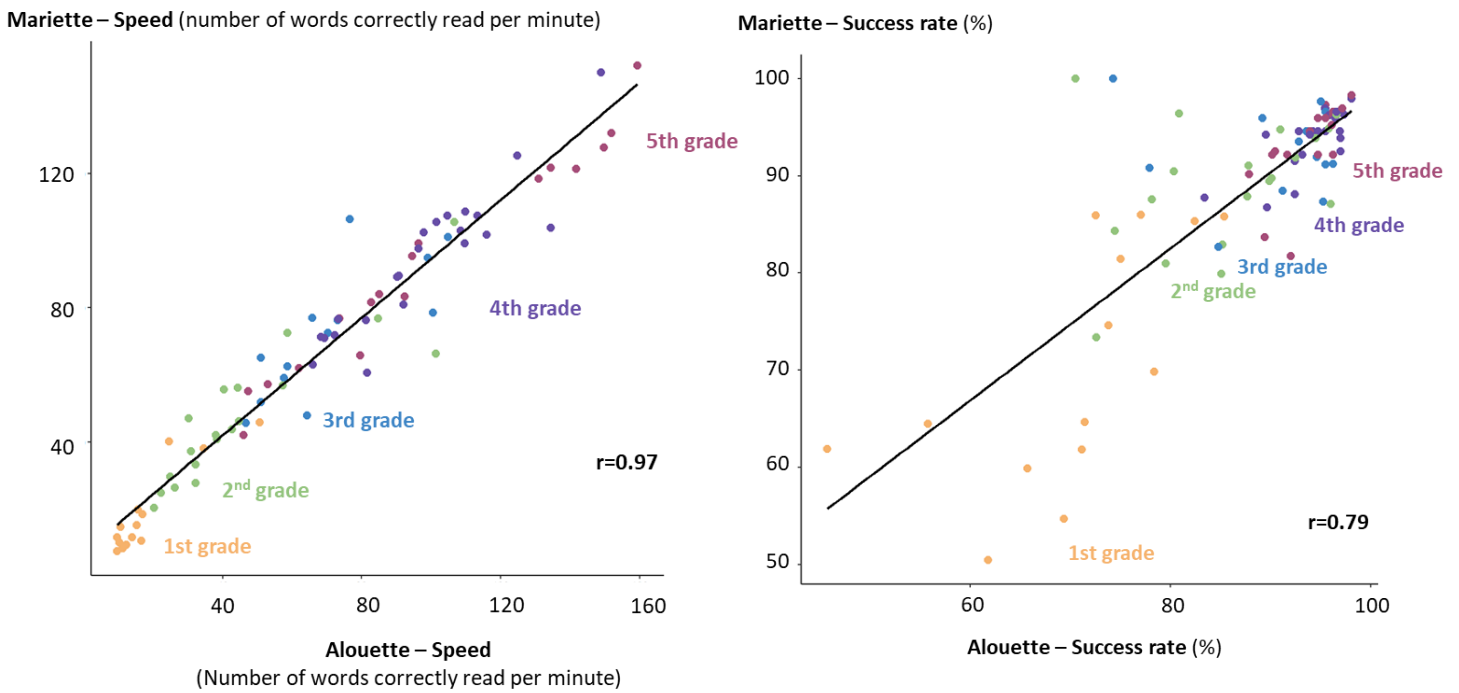


**Figure 5.** Evolution of the percentage of self-corrections with schooling

### Validity of the Mariette

Finally, in order to measure the empirical validity of our test, we correlated the speed and success rate scores on the Alouette and Mariette tests for 20% of the pupils assessed in March. We chose the Alouette test because it is one of the tests that most closely resembles the one we developed, since it is also meaningless, includes a similar number of words (265/294) and has a comparable completion time (3 minutes/5 minutes).

The results, graphically depicted in Figure 6, showed a massive correlation between the two tests, both in terms of speed ( $r=0.97$ ,  $p<0.001$ ) and success rate ( $r=0.79$ ,  $p<0.001$ ). So the Mariette measures very similar performance to the Alouette, while allowing sophisticated error analysis.



**Figure 6.** Correlations between the Alouette and the Mariette

## DISCUSSION

The change in reading speed observed here between grades is similar to that observed in other longitudinal and cross-sectional studies, with a quasi-linear increase as a function of schooling for the youngest pupils, including during 1<sup>st</sup> and 2<sup>nd</sup> grade, and a plateau that begins to be reached towards the end of primary school (Al Otaiba et al., 2009; Biemiller, 1977; Lervag & Hulme, 2009; Nese et al., 2013). A great deal of research has shown that reading speed, also known as fluency, is a good reflection of a child's reading ability. A marker of automaticity in reading, this measure is highly correlated with comprehension (Fuchs et al., 2001; Hudson et al., 2005; Lee & Chen, 2019). Indeed, children who reads automatically will be able to pay more attention to the meaning of the words they read and make their working memory more available.

Reading speed increases monotonically as schooling progresses, and this growth is significant at every stage, except between 4<sup>th</sup> and 5<sup>th</sup> grade, but including between the end of 1<sup>st</sup> grade and the beginning of 2<sup>nd</sup> grade. However, for the same number of months (6 months separate the measurements at the

start and end of 1<sup>st</sup> grade, all the same as the measurements at the end of 1<sup>st</sup> grade and the start of 2<sup>nd</sup> grade), this growth is smaller between school years than during the year. Given that pupils do not receive explicit teaching during the vacations that separate 1<sup>st</sup> and 2<sup>nd</sup> grade, this result is not surprising in itself, but could suggest that a fraction of pupils benefit from 'self-learning' (Share, 1995) : once decoding is in place, their reading continues to help them make progress, because the phonological pathway enables them to decode the pronunciation of new words, and therefore to recognize them if they belong to their oral vocabulary, and thus provides the lexical pathway with data to learn about the correlation between spelling, pronunciation and meaning. It should be remembered, however, that there is also a negative effect of the holidays, recently quantified by the DEPP: during the summer between 1<sup>st</sup> and 2<sup>nd</sup> grade, many children make no further progress, especially those from poorer backgrounds (Andreu et al., 2023). Finally, it should be remembered that the data presented here are cross-sectional and not longitudinal. As a result, it is not the same children who contribute to the data for 1<sup>st</sup> and 2<sup>nd</sup> grade, and we cannot rule out the possibility that the recent measures introduced to improve reading instruction are beginning to bear fruit. These three factors probably contribute to the differences observed between the end of 1<sup>st</sup> grade and the beginning of 2<sup>nd</sup> grade, although it is not possible in our experience to separate them. These effects, however, remain minimal.

The success rate is also a parameter that is highly correlated with reading comprehension, which is interesting to consider in more opaque languages such as French (Ziegler et al., 2010) for which it takes time to master all the rules of reading. Success rates in more opaque languages often take several months to approach 100% (Seymour et al., 2003; Vaessen et al., 2010; Ziegler & Goswami, 2005), which is also what we observed here. The plateau of around 95% success rate observed from 4<sup>th</sup> grade onwards is consistent with the automaticity of reading acquired by pupils at the end of primary school. A percentage of errors of around 5% persists, due in particular to certain rare irregular words ("dolmen") and also to the presence of pseudowords close to frequent French words ("truche", "arbi", "oviles", "mantou").

When we looked at error rates for different types of stimuli, we observed the presence of significant lexicality and regularity effects from 1<sup>st</sup> to 5<sup>th</sup> grade. This is consistent with the literature, which shows that the first lexicality effects appear only a few months after the start of reading (Zoccolotti et al., 2009) and reports the existence of a regularity effect already in 3<sup>rd</sup> graders (Schmalz et al., 2013). These two studies mention the modulation of these effects with frequency, since the lexicality effect first appears between frequent regular words and pseudowords before extending later to rarer regular words, and the regularity effect is present only on low-frequency words, as in adults (Paap & Noel, 1991; Seidenberg et al., 1984), a modulation that unfortunately we can't see here, because the number of stimuli is too small.

For transposable words, we observed with the Mariette an overall decrease in the error rate on these words and an error rate higher than the error rate on regular non-transposable words. These observations are compatible with the results of Paterson et al. (Paterson et al., 2015), which show that the error rate on transposable words is higher than that on non-transposable words in adults and in 2<sup>nd</sup>-4<sup>th</sup> graders, as well as a decrease in the error rate on transposable words between 2<sup>nd</sup>-4<sup>th</sup> graders and adults. However, we did not observe a bell curve with a maximum reached in 3<sup>rd</sup> grade as Grainger et al. (2012) but an almost monotonic decrease, probably because we didn't have enough items to focus solely on transposable pseudowords, which were the subject of their study.

Our results on reading fluency and on the different types of error observed are therefore consistent with those described in the literature. The evolution of the precise error rate for each type of error then provided us with benchmarks for the relative frequency of these errors within a grade and their number, which we could interpret. For example, it is perfectly normal for a child in 1<sup>st</sup> grade to pronounce silent endings, make mistakes with digrams or not yet master contextual rules. As French is a rather opaque language, it takes at least a year of learning to read for all these rules to be mastered (Seymour et al., 2003; Vaessen et al., 2010; Ziegler & Goswami, 2005). As we can see here, mastery of these contextual rules continues to pose problems for older pupils, which is in line with other studies

that have shown that pupils are still making many errors of this type when they enter 6th grade (Lubineau et al., 2023). Confusions over functional words, which are present in large proportions in the oldest pupils, should be regarded as normal errors, a sign of automaticity and speed of reading, since the eye hardly ever looks at these short, very frequent words (Rayner & Duffy, 1986; Rayner & McConkie, 1976). Our results also showed that other types of error diminished over time. This is why their existence in some pupils, even in small numbers, should be a cause for alarm. This is particularly true of the mispronunciation of digrams and the pronunciation of silent final letters.

The Mariette is part of a wider range of tests that should be used together to formulate a diagnosis. Its standards help to guide the interpretation of a child's errors, and to determine whether they can be considered normal or whether, for certain types of error, their high number could be a sign of an underlying disorder, in which case further tests will need to be carried out. For example, if a pupil makes more letter position errors than average, but makes very few attentional and visual errors, there is a good chance that he or she has letter position dyslexia, but this initial impression will have to be confirmed later by taking the Malabi.

The work presented here is still in progress and some adjustments still need to be made in order to establish precise standards for each type of error and to ensure that these standards are meaningful. Our main obstacle concerns the fact that not all the students read the same number of words. This raises the question of how many words a student needs to have read for it to be relevant to look at the types of errors they have made. We also need to confirm the validity of our approach by testing the Mariette and Malabi with dyslexic and normal-reading pupils to establish the correlation between these two exercises.

## **CONCLUSION**

Our measurements show that the Mariette is a good test for assessing reading speed and accuracy in primary school pupils. What makes it so useful is the detail we were able to quantify about the exact nature of the error made by the child. The changes we have observed in the exact types of errors made



by pupils according to their level will, for the first time, provide practitioners and teachers with benchmarks for the different errors made by their pupils, so that they can help them better. However, the test does not provide a sufficiently accurate diagnosis to determine the exact type of dyslexia a child may have. It is a first step towards more detailed tests that give practitioners an indication of the child's potential deficit(s), thus pointing them in the direction of further investigation.

This test, developed in partnership with CERENE schools and diagnostic centre, will shortly be available for a few euros via a manual and a dedicated application, which will include the text, standards for speed, success rate, fatigability, self-corrections and types of error, as well as scoring help.

**CHAPTER 3 – DOES WORD FLICKERING IMPROVE  
READING? NEGATIVE EVIDENCE FROM FOUR  
EXPERIMENTS USING LOW AND HIGH FREQUENCIES**

## **ABSTRACT**

Does word flickering facilitate reading? Despite a lack of scientific evidence, flickering glasses and lamps for dyslexia are being marketed in various countries. We conducted four experiments to assess their efficacy. Two experiments involved a computerized lexical decision task with constant display or low-frequency flickering (10 or 15 Hz). Among 375 regular adult readers, flicker noticeably slowed down word recognition, while slightly biasing the decision towards pseudowords. No significant effect was observed in 20 dyslexic children. In 22 dyslexic children, we also evaluated the impact of the Lexilight® lamp and Lexilens® glasses, which operate at higher frequencies, on reading fluency, letter identification and mirror letter processing. No detectable impact was observed. Lastly, in two participants who claimed to benefit from flickering glasses, we orthogonally manipulated whether the glasses were actually on, and whether the participant thought they were on. Only a small placebo effect was noted in one participant. Our findings starkly contrast with marketing claims that these tools can help 90% of dyslexics, and emphasize the role of rigorous scientific research in empowering dyslexic individuals to make informed decisions.

## **INTRODUCTION**

Dyslexia is a neurodevelopmental disorder characterised by important difficulties in reading acquisition in the presence of normal intelligence and access to education. It is estimated that 3-12% of children are affected by this disorder, depending on language and dyslexia definition (Di Folco et al., 2022; Lindgren et al., 1985). Many different directions are pursued in dyslexia research, including the existence of subtypes, their behavioural characterization, and their cognitive, circuit-level, neuronal and genetic mechanisms. Here, leaving those questions aside, we concentrate on one issue: is it possible to facilitate reading for children with dyslexia by manipulating their reading experience? Assessing this issue scientifically is all the more important that many companies are quick to market products for dyslexic populations that are often labelled as life changing, generally without any supportive evidence.

Some of the tools available on the market offer a variety of options to modify the layout of texts. For instance, electronic book readers offer the option of enlarging the font size, spacing the lines further apart, changing the background colour of the page or changing the font to a special font such as Dyslexia, OpenDys or EasyReader. While these technologies claim to facilitate reading, increasing font size and character spacing are the only options that have so far proven to be effective. A study by O'Brien et al. found that reading speed improved with font size in all students, then reached a plateau, with the font size at which this plateau was reached being slightly larger for dyslexics (O'Brien et al., 2005). This facilitating effect of font size was confirmed by another study, conducted by Rello and Baeza-Yates, which shows that the reading of dyslexic pupils is improved by using a font size of 22 or 26 points, compared with a font of 14 points (Rello & Baeza-Yates, 2017). While it may simply indicate that many dyslexics, similar to beginner readers, have not yet adapted to small print, it does offer a simple way to help them.

Increasing the spacing between characters is also an effective parameter (Zorzi et al., 2012), that could even be more effective than increasing the font size (Katzir et al., 2013). Spacing letters a few extra

percent apart has been reported to increase reading speed, reduce errors and facilitate comprehension (Łuniewska et al., 2022; Stagg & Kiss, 2021; Duranovic et al., 2018; Rello & Baeza-Yates, 2017; Hakvoort et al., 2017; Sjoblom et al., 2016; Zorzi et al., 2012). To be optimal, it has to be combined with an increase in the spacing between words (Galliussi et al., 2020).

A study from Rello and Baeza-Yates found no facilitating effect from the use of 1.4 line spacing, a coloured or grey background with black or white writing, or the use of a specific font (Rello & Baeza-Yates, 2017). In another study, no effect of the “dyslexia font” OpenDyslexic was observed on reading rate and accuracy (Wery & Diliberto, 2017). The effectiveness of specific fonts which is occasionally reported (Bachmann & Mengheri, 2018) fades once controlling for the spacing between characters, which is greater in specific fonts (Joseph & Powell, 2022; Galliussi et al., 2020; Marinus et al., 2016), thus suggesting that character spacing, rather than the font itself, is the most impactful variable.

Regarding the effect of colors, Humphreys and Mayall (2001) and Friedmann and Rahamim (2014) reported that colouring each letter using a different colour did not improve their dyslexic participants' results, and in some cases even worsen them, compared to baseline. Other studies, this time involving groups of dyslexic children, yield the same conclusions (Koornneef & Kraal, 2022; Pinna & Deiana, 2018).

Recently, a new idea has emerged among manufacturers: flickering words using either stroboscopic light or flickering glasses. Either the light emitted by the lamp flickers at a very high, almost imperceptible frequency (from 60Hz to 120Hz), or the glasses' lenses darken and light up, also at a very high frequency (from 70Hz to 90Hz). The scientific rationale behind this idea seems extremely thin. It stems from a study by Le Floch and Ropars (2017), published in this journal, who claimed that dyslexia is caused by a retinal anomaly leading to the formation of illusory mirror images and resulting in the absence of a dominant eye, that could be remedied by high frequency flickering. The logic of this study is highly debatable: dyslexia was never properly tested, as no reading scores were provided; statistics were flimsy; a retinal anomaly, if it was properly documented, would not explain the dissociations

observed in dyslexia, for instance between number and letter reading (Dotan & Friedmann, 2019; Friedmann, Dotan, et al., 2010); why flickering would bypass it remains unclear; and finally, to the best of our knowledge, the results of this study have never been replicated.

Thus, it would seem easy to dismiss flickering as an eccentric proposal, were it not for several possible counterarguments. First, manufacturers seem successful in selling their products. Second, physiological recordings show that even subjectively invisible flicker frequencies can induce rhythmic neural activity in lateral geniculate and primary visual cortex (Gur & Snodderly, 1997; Krolak-Salmon et al., 2003). Third, prolonged adaptation to fast flickering visual noise can improve acuity in fine visual recognition tasks, including word recognition in a small font (Arnold et al., 2016; J. Lee & Chong, 2021). Finally, and most relevantly, a few studies have described adults with mirroring reading disorders who were helped by flickering (McCloskey et al., 1995; McCloskey & Rapp, 2000a; Pflugshaupt et al., 2007; Vannuscorps et al., 2022). In particular, a single case of developmental dyslexia, documented in great detail by McCloskey and collaborators in a series of articles, presented with a severe confusion of right and left, frequently copied figures in mirror image and, when reading, often mirrored letters, for instance reading lamp as lamb (McCloskey et al., 1995; McCloskey & Rapp, 2000a). Her word reading errors arose at a visual level prior to semantic access. Remarkably, her mirror effects vanished, and reading became almost perfect, when stimulus exposure time was low (<100ms) or under low-frequency flickering (10Hz). Those factors led to an abrupt transition from a very low error rate (0.5% in reading a word list) to a much higher error rate (25%). McCloskey et al. tentatively interpreted this flickering effect as a reflection of the subdivision of the visual system into a transient subsystem specialized for processing rapidly changing visual stimuli, the magnocellular pathway, and a sustained subsystem more sensitive to static or longer-duration stimuli, the parvocellular pathway. Both pathways link the retina to the visual cortex through ganglion and bipolar cells (Masri et al., 2020). McCloskey patient's behaviour might have arisen from an impairment in the parvocellular pathway, which would have been short-circuited by flickering, thus activating only the magnocellular pathway, supposed to be intact. Since these pathways start in the retina, a putative impairment of the

parvocellular pathway could perhaps be related to a different organisation of retinal cells, thus establishing a tentative connection with le Floch and Ropars (2017)'s paper.

Another patient described by Pflugshaupt et al. (2007) acquired mirror writing and reading following brain damage. Again, her reading came back to normal under low frequency flickering (10Hz) with an abrupt transition between presentation durations of 100ms and 200ms, where her performances suddenly worsened. Finally, the patient described by Vannuscorps et al (2022) perceived high-contrast shapes as if they had rotated 90° or 180°, or mirrored the initial shape. Her orientation difficulties disappeared almost completely when the stimulus flickered at 5.7Hz. Note that these results were obtained at low frequencies, quite far from the frequencies mentioned by Le Floch and Ropars, which were higher than 70 Hz. Nevertheless, flickering clearly helped the patients, thus begging the question of whether it could benefit other dyslexia patients or the general population. In this study, we therefore aimed to better understand if low or high frequency flickering could facilitate reading for normal readers and dyslexics. Figure 1 summarizes our approach. First, we studied the impact of low-frequency flickering, similar to McCloskey et al., and Pflugshaupt et al., on reading performance in normal adults and dyslexic children. Next, we turned to the impact of high frequency flickering. For this, we used a more natural setting (reading on paper) and the lamp and glasses described above. We tested a flickering frequency of ~80Hz, first on a group of dyslexic children unfamiliar with these devices, and second on two patients, one adult and one child, who both claimed to be helped by the glasses on a daily basis.

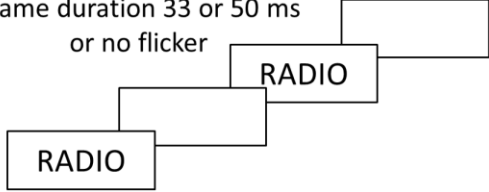

Participants	Task	Frequency	Types of displays
<b>Experiment 1</b> 375 normal adult readers	Lexical decision	10 or 15 Hz	<b>Reading on computer screen</b> Frame duration 33 or 50 ms or no flicker 
<b>Experiment 2</b> 20 dyslexic children	Lexical decision	10 or 15 Hz	
<b>Experiment 3</b> 22 dyslexic children	Letter identification Single words reading Text reading	80 Hz	<b>Reading on paper</b> Glasses                      Lamp  Or natural light
<b>Experiment 4</b> 2 dyslexics, one adult and one child, claiming to be helped by the glasses	Single words reading Text reading Reading comprehension	Preferred frequency (FAP: 87Hz CT: 82Hz)	

Figure 1. Logic of our successive experiments

## EXPERIMENT 1. IMPACT OF LOW-FREQUENCY FLICKERING ON NORMAL ADULT READERS

### Method

#### *Participants*

Participants were recruited via Twitter. The study was conducted on-line on a computer or a touch screen, during the Covid19 epidemic. Participants were informed that they could leave the task at any time and that in this case their data would not be retained. 778 adults participated (543 females, 226 males, age group breakdown: 18-24 years old, 38 subjects; 25-40, 239 subjects; 41-60, 442 subjects; >60, 59 subjects).



### **Lexical Decision Task**

To measure the effect of low-frequency flickering on reading, we used a lexical decision task. The stimuli were randomly drawn for each student from the database extracted from Lexique 3.83 (New et al., 2004) described by Lubineau et al. (2023). Words varied in length (4 to 8 letters) and fell into four different frequency bands: very frequent, frequent, rare and very rare (see Table 1 for details and examples). Pseudowords were also between 4 and 8 letters long, and divided into 6 categories, according to the nature of the trap they presented (see Table 2 for details and examples). These pseudowords categories were matched and came in pairs as followed: orthographic traps and word approximations, transpositions and double substitutions and mirror and single substitutions.

Frequency category	Examples	
	word	translation
Very frequent (greater than 100 per million)	beau	nice
	message	message
Frequent (40 to 100 per million)	usine	factory
	étudier	study
Rare (10 to 40 per million)	carnet	booklet
	éprouver	experience
Very rare (3 to 10 per million)	cerf	deer
	abolir	abolish

**Table 1.** Characteristics and examples of single-word stimuli for experiments 1-2

Pseudoword traps	Description	Examples		
		pseudoword	associated word	translation
Orthographic traps	Created from words by manually introducing orthographic mistakes. They can be read as words if the participant does not correctly master the grapho-syntactic rules of French.	bage	bague	ring
		inciet	inquiet	worried
Word approximations	Control for orthographic traps. Assembly of trigrams according to a markov procedure to ensure a probability of occurrence of syllables similar to that of French.	atio		
		ouvoi		
Transpositions	Created from words by inverting two adjacent consonants or vowels	ceil	ciel	sky
		pafrois	parfois	sometimes

Double substitutions	Control for transposition created by substituting the same consonants or vowels by two others.	cuol pansois	ciel parfois	sky sometimes
Mirror substitutions	Created from words by applying the following rules:  $b \rightarrow d / d \rightarrow b / p \rightarrow q / q \rightarrow p$	qièce dateau	pièce bateau	room boat
Single substitutions	Control for mirror substitutions created by applying the following rules:  $b \rightarrow f / d \rightarrow t / p \rightarrow g / q \rightarrow j$	gièce fateau	pièce bateau	room boat

**Table 2.** Characteristics and examples of pseudoword stimuli for experiments 1-2

This lexical decision comprised 360 stimuli (180 words and 180 pseudowords), each randomly presented in one of three conditions: a continuous display, a flickering display at 10Hz or a flickering display at 15Hz. These flickers were such that the stimulus was displayed in the first half of the period and replaced by an empty screen in the second half, exactly as in previous publications (McCloskey & Rapp, 2000a; Pflugshaupt et al., 2007). For each of these trials, we collected accuracy and response time.

### **Data analysis**

Because the experiment was run on-line, some of the trials were presented at a duration that departed from the desired regular flickering. We excluded participants for whom more than 20% of the stimuli had the wrong timing. This resulted in a smaller sample of 375 participants (246 women, 123 men, age group breakdown: 18-24, 16 subjects; 25-40, 120 subjects; 41-60, 208 subjects; >60, 31 subjects). This subsample is equivalent to the original one in terms of sex,  $\chi^2(1)=1.64$ ,  $p=0.20$ , and age,  $\chi^2(3)=0.57$ ,  $p=0.90$ . We further excluded trials with response times below 200ms or inappropriate timing (2.4% of trials). The remaining 131,714 trials were then analysed with mixed-effects models, using a procedure similar to that described by Lubineau et al (2023), incorporating display condition (3 levels, continuous display, flickering at 10Hz or flickering at 15Hz), lexicality (word, pseudoword), length (4-8 letters),

word frequency (a numeric variable encoding the frequency category of the word) pseudoword-type (which we analysed by two-by-two comparisons for matched pairs) as fixed effects. We run three successive models, with the following structure

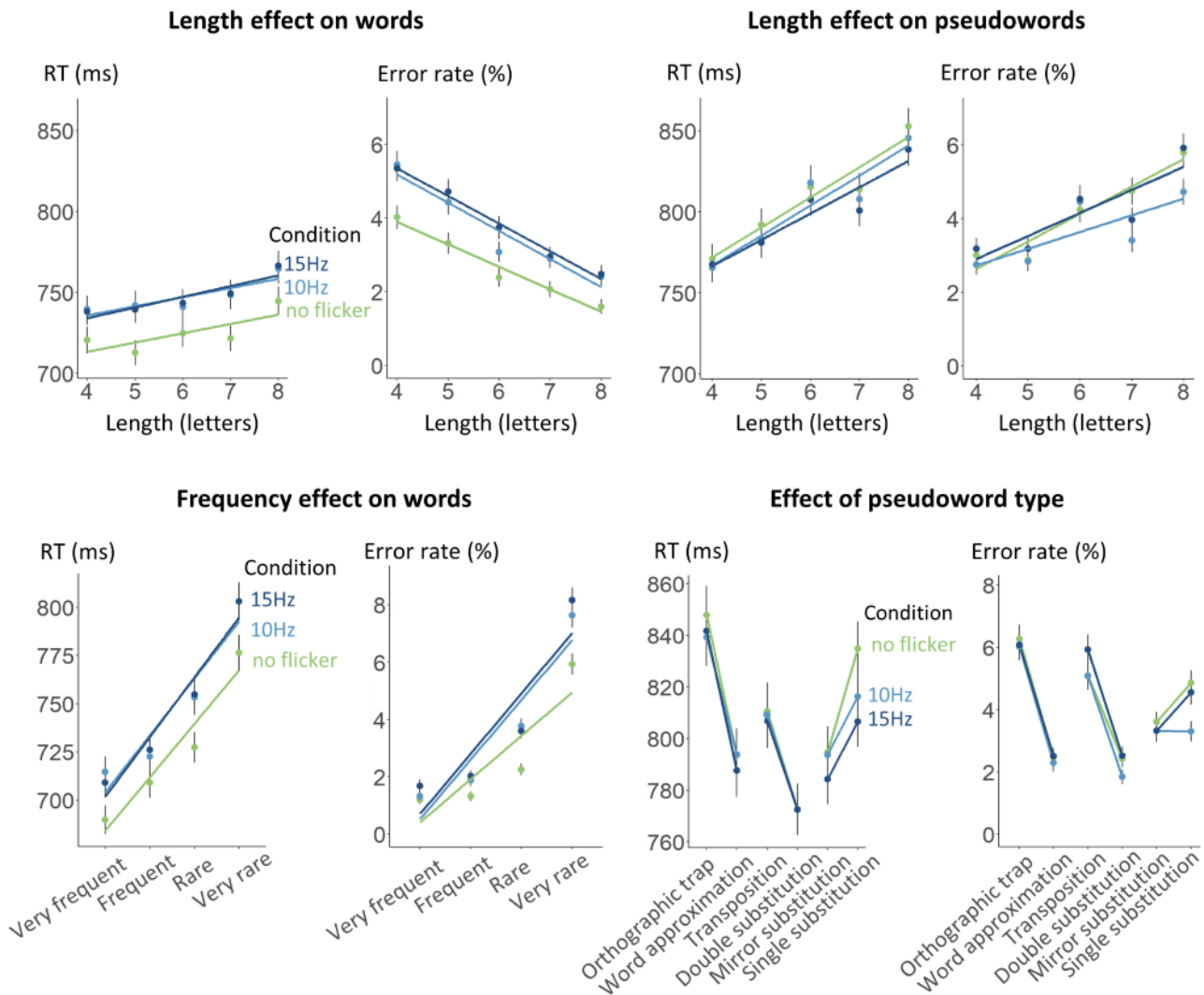
$$dv \sim X_1 * X_2 * \dots * X_n + (1|subject) + (1|stimulus)$$

with the dependent variable being either response time (RT) or error rate (ER).  $X_i$  represents the combination of our interacting fixed effects. We used subject and item as random effects. All response time analyses were performed on correct responses only. Error data were submitted to the same item-based models, using logistic mixed effect models. Whenever a frequentist test evaluated the effect of flickering displays, we also used Bayesian statistics to evaluate the weight of evidence for or against the hypothesis that it had an effect on performance. We used the BayesFactor R package to compute Bayesian mixed effect models and obtain Bayes factors for each effect. We used the models described above, except that we only put participant as a random variable in order to maintain a reasonable computation time. With our conventions, a Bayes factor (BF) between 3 and 10 offers substantial evidence that flickering has an effect, while  $BF > 10$  is strong evidence. In the opposite direction, a BF between 0.33 and 0.1 offers substantial evidence in favour of the null hypothesis that flickering has no effect, and  $BF < 0.1$  is strong evidence.

## Results

We start by summarizing the effects of lexicality, length, frequency, and type of pseudoword. Figure 2 graphically depicts these effects as a function of display condition (continuous display, flickering at 10Hz or flickering at 15Hz). The results tightly replicated our previous findings with the same task but without flickering (Lubineau et al., 2023). A lexicality effect was found only on response times (RT) with pseudowords taking longer to classify than words (RT:  $F(1,3016.0)=334.63$ ,  $p < 0.001$ ,  $BF > 100$ ; ER:  $\chi^2(1)=0.25$ ,  $p=0.62$ ,  $BF > 100$ ). The results also confirmed the significance of the length effect (RT:  $F(1,2810.8)=106.08$ ,  $p < 0.001$ ,  $BF > 100$ ; ER:  $\chi^2(1)=5.89$ ,  $p=0.015$ ,  $BF=0.042$ ) and its interaction with

lexicality (RT:  $F(1,2810.8)=9.82$ ,  $p=0.002$ ,  $BF>100$ ; ER:  $\chi^2(1)=53$ ,  $p<0.001$ ,  $BF>100$ ), showing a larger length effect for pseudowords than for words.



**Figure 2.** Deleterious effect of low-frequency flickering (10Hz and 15Hz) on lexical decision in normal adults. Each point represents the mean RT or error rate as function of word length, frequency or pseudoword type and condition. Error bars represent one standard error of the mean. The slopes are the linear regression associated with the points.

Within words, the frequency effect was significant, with higher RT and error rates for lower frequency words (RT:  $F(1,1796.1)=611.82$ ,  $p<0.001$ ,  $BF>100$ ; ER:  $\chi^2(1)=344.68$ ,  $p<0.001$ ,  $BF>100$ ). Finally, within pseudowords, participants were slower and less accurate at spotting orthographic traps, compared to word approximations (RT:  $F(1,327.18)=28.90$ ,  $p<0.001$ ,  $BF>100$ ; ER:  $\chi^2(1)=29.91$ ,  $p<0.001$ ,  $BF>100$ ) and at spotting transpositions compared to double substitutions (RT:  $F(1,368.04)=25.30$ ,  $p<0.001$ ,  $BF>100$ ; ER:  $\chi^2(1)=48.10$ ,  $p<0.001$ ,  $BF>100$ ). We replicated our previous observation of a paradoxical effect of

mirror substitutions, which were faster to classify than single substitutions,  $F(1,300.76)=7.71$ ,  $p=0.006$ ,  $BF>100$ , although their error rate did not differ,  $\chi^2(1)=2.56$ ,  $p=0.11$ ,  $BF=2.7$ . We explained this effect by the fact that some mirror-letter substitutions involving letter “q” violated the orthographic statistics of French and such violations facilitated the rejection of mirror substitution pseudowords (see Lubineau et al 2023).

Crucially for our current purposes, flickering affected some of these observations. Flickering at 10Hz or 15Hz, relative to a continuous display, slowed down words responses (RT:  $F(2,62001.1)=61.68$ ,  $p<0.001$ ,  $BF>100$ ) and made them moderately more error-prone (ER:  $\chi^2(2)=57.44$ ,  $p<0.001$ ,  $BF=0.011$ ), while slightly facilitating responses to pseudowords (RT:  $F(2,62134.7)=6.36$ ,  $p=0.0017$ ,  $BF=0.047$ ; ER:  $\chi^2(2)=9.45$ ,  $p=0.009$ ,  $BF=0.023$ ). Those effects led to a strong interaction between display condition and lexicality (RT:  $F(2,124763.1)=49.11$ ,  $p<0.001$ ,  $BF>100$ ; ER:  $\chi^2(2)=42.49$ ,  $p<0.001$ ,  $BF>100$ ), which can be described as a bias towards classifying items as pseudowords whenever they were flickering. Flickering did not impact the length effect, as the interaction between length and condition was not significant (RT:  $F(2, 124378.0)=0.24$ ,  $p=0.79$ ,  $BF<0.01$ ; ER:  $\chi^2(2)=0.91$ ,  $p=0.64$ ,  $BF<0.01$ ). There was no significant interaction of flickering condition and word frequency (RT:  $F(2,62456.8)=1.64$ ,  $p=0.19$ ,  $BF<0.01$ ; ER:  $\chi^2(2)=0.59$ ,  $p=0.74$ ,  $BF=3.1$ ).

Finally, concerning specific comparisons between pseudowords, there was no effect of flickering on all pseudowords comparisons (orthographic traps versus word approximation: RT:  $F(2,20418.12)=0.80$ ,  $p=0.45$ ,  $BF<0.01$ ; ER:  $\chi^2(2)=0.47$ ,  $p=0.79$ ,  $BF<0.01$ ; transposition versus double substitutions: RT:  $F(2,20501.04)=0.179$ ,  $p=0.84$ ,  $BF<0.01$ ; ER:  $\chi^2(2)=1.93$ ,  $p=0.38$ ,  $BF<0.01$ ; mirror versus single substitutions: RT:  $F(2,20471.02)=1.78$ ,  $p=0.17$ ,  $BF<0.01$ ; ER:  $\chi^2(2)=4.47$ ,  $p=0.11$ ,  $BF=0.018$ ).

Thus, flickering displays did not enhance reading in normal adults. Instead, it impaired lexical decision for written words and biased participants towards the pseudoword response. While this result could be described as a performance improvement for pseudowords, the simplest explanation is that flickering displays looked slightly abnormal to participants and therefore biased them towards the

pseudoword response. However, it remains possible that flickering selectively facilitates reading in dyslexics. In experiment 2, therefore, we used exactly the same lexical decision in dyslexic students.

## **EXPERIMENT 2 – IMPACT OF LOW FREQUENCY FLICKERING ON**

### **DYSLEXIC CHILDREN**

#### **Method**

##### ***Participants***

Our participants were all dyslexic students coming from the CERENE schools, specialized for students with learning disabilities and with normal intelligence. Class sizes and teaching methods are adapted to the needs of these students.

29 CERENE students, from 6<sup>th</sup> to 8<sup>th</sup> grade, took part in this study, all of them diagnosed as dyslexic by a professional speech therapist. To confirm the diagnosis, they first took the Alouette test, a test used in France to test for dyslexia. Only those students whose Alouette scores were more than 1.5 standard deviations below the mean in speed or error rate were retained. 22 students met this criterion. 20 students finally completed the task, as two were absent on the day of the tests.

##### ***Procedure***

The procedure was the same as in Experiment 1, except that the students did not take the task online but during their school time. Parents were informed of the experiment by mail beforehand, and the participation of their child in the study was subject to their approval. Each participant performed the test individually in a quiet room.

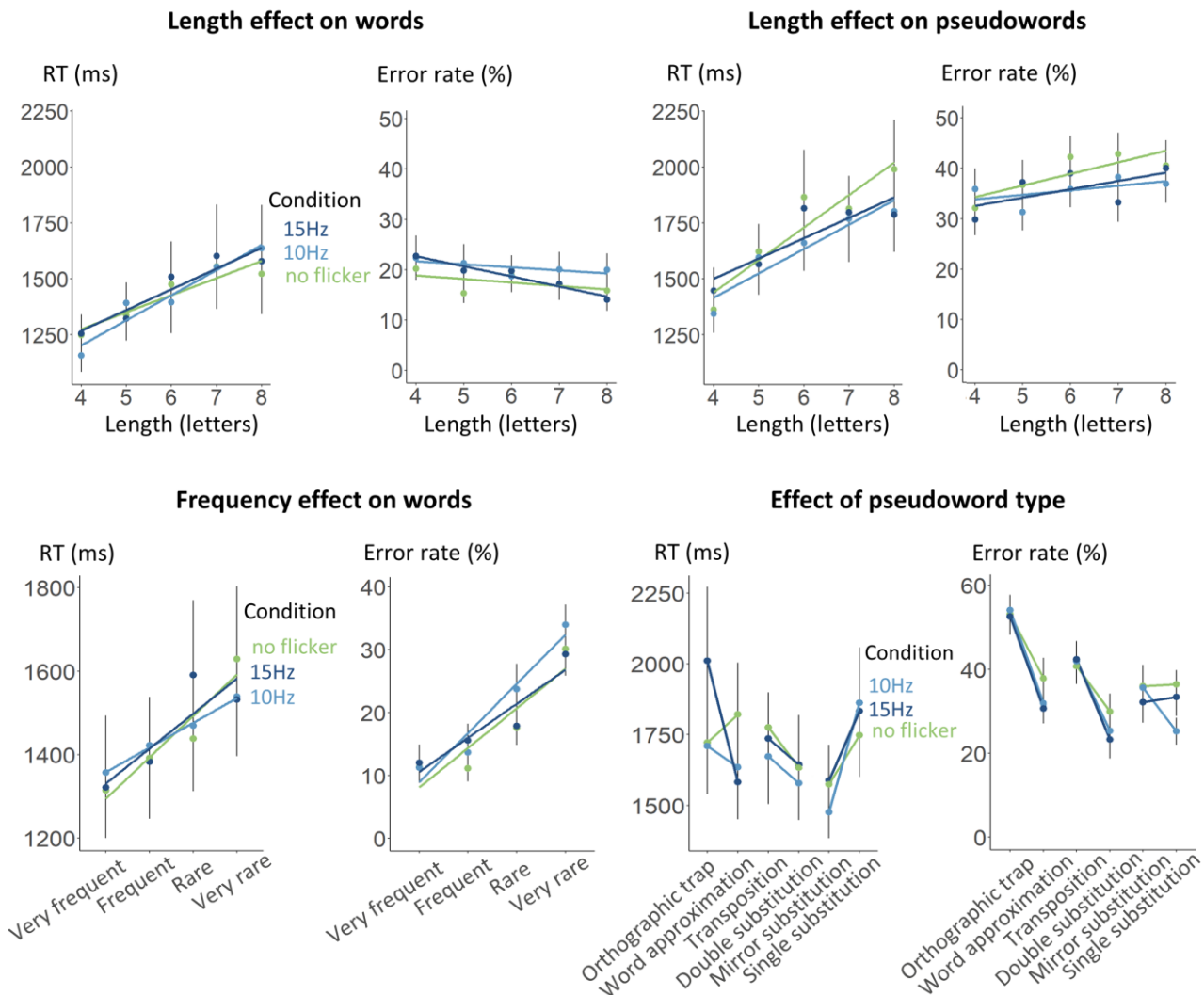
##### ***Data analysis***

Data cleaning followed the same rules as in the previous experiment. Trials with a RT of less than 200ms and trials with imperfect timing (4.1% of trials) were excluded. RTs that fell 3 standard deviations or

more above the subject's mean were also excluded. Finally, one student was excluded because his performance did not differ from chance ( $\chi^2(1)=0.04$ ,  $p=0.53$ ). The remaining 6,905 trials were then analysed with mixed-effects models, using a procedure similar to that described in Experiment 1.

## Results

Results are shown in Figure 3. Mixed-effects models show that RT and error rate were higher for pseudowords than for words (RT:  $F(1,1693.8)=124.66$ ,  $p<0.001$ ,  $BF>100$ ; ER:  $\chi^2(1)=194.47$ ,  $p<0.001$ ,  $BF>100$ ). A large effect of length, typical for dyslexic readers, affected RT but not error rates (RT:  $F(1,1290.5)=176.93$ ,  $p<0.001$ ,  $BF>100$ ; ER:  $\chi^2(1)=0.52$ ,  $p=0.47$ ,  $BF=0.11$ ). There was also a significant effect of frequency (RT:  $F(1,1410.4)=43.83$ ,  $p<0.001$ ,  $BF>100$ ; ER:  $\chi^2(1)=108.63$ ,  $p<0.001$ ,  $BF>100$ ). Finally, the patterns of differences between pseudowords replicated those observed by Lubineau et al. with their least fluent students. Orthographic traps were more error prone than word approximations,  $\chi^2(1)=32.67$ ,  $p<0.001$ ,  $BF>100$ ; transpositions were more error prone than double substitutions,  $\chi^2(1)=29.42$ ,  $p<0.001$ ,  $BF>100$ ; and, we observed faster responses to mirror substitutions than to single-letter substitutions,  $F(1,218.93)=14.76$ ,  $p<0.001$ ,  $BF>100$ .



**Figure 3.** Lack of impact of low-frequency flickering (10Hz and 15Hz) on lexical decision in 20 dyslexic children. Each point represents the mean RT or error rate as function of word length, frequency or pseudoword type and condition. Error bars represent one standard error of the mean. The slopes are the linear regression associated with the points.

Crucially, flickering had very little impact on these observations. The main effect of display condition was not significant in any of the analysis performed, the interaction of condition X lexicality did not reach significance either (RT:  $F(2,4714.1)=0.86$ ,  $p=0.42$ ,  $BF<0.01$ ; ER:  $\chi^2(2)=5.74$ ,  $p=0.057$ ,  $BF=0.070$ ), nor did the interaction between length and condition (RT:  $F(2,4699.7)=1.82$ ,  $p=0.16$ ,  $BF<0.01$ ; ER:  $\chi^2(2)=0.85$ ,  $p=0.65$ ,  $BF<0.01$ ). There was no interaction with frequency (RT:  $F(2,2643.5)=0.68$ ,  $p=0.51$ ,  $BF=0.012$ ; ER:  $\chi^2(2)=2.36$ ,  $p=0.31$ ,  $BF=0.023$ ).

Only a single, barely significant interaction was found on RTs to pseudowords, when comparing orthographic traps and word approximations ( $F(2,587.18)=3.19$ ,  $p=0.042$ ,  $BF=0.76$ ). This effect was



unsupported by Bayesian comparison and would not have survived a correction for multiple comparisons. It suggested that at a frequency of 15Hz only, dyslexic students were slower to detect orthographic traps compared to words approximations. Even if this effect was deemed significant, it would correspond to an impairment rather than a facilitation by flickering.

Thus, we found that flickering had no facilitating effect in 20 dyslexic students. Contrary to what we observed with adults, we did not find any slowdown in the detection of words when they were flickering – but crucially, there was no facilitation either.

### **EXPERIMENT 3 – IMPACT OF HIGH FREQUENCY FLICKERING ON A GROUP OF DYSLEXIC STUDENTS**

Experiments 1 and 2 tested flicker frequencies of 10 and 15 Hz, based on their effectiveness in previous single-case studies (McCloskey & Rapp, 2000a; Pflugshaupt et al., 2007). However, commercially available devices for dyslexic individuals use much faster frequencies and claim to aid reading on paper. To come as close as possible to the conditions that are claimed to be effective by the manufacturers, we next tested dyslexic students using those devices, using purely paper-based tests.

#### **Method**

##### ***Participants***

35 dyslexic students from CERENE participated in this second study, 20 of whom had already taken part in the previous one. As in Experiment 2, all of them first took the Alouette test to confirm the presence of dyslexia. The results showed that 28 of these students were below standard in speed or error rate. As the various testing sessions described below took place over several weeks, only 22 pupils, from 4<sup>th</sup> to 8<sup>th</sup> grade, were able to take part in all sessions.

### ***Test procedures***

To study the impact of high-frequency flickering on dyslexic students' reading, we used the Lexilens<sup>®</sup> glasses and the Lexilight<sup>®</sup> lamp, unfamiliar to the students. The light emitted by the Lexilight<sup>®</sup> lamp flickers at an almost imperceptible frequency, between 60Hz and 120Hz, and can be adjusted among 5 different frequencies. The Lexilens<sup>®</sup> glasses use electronic lenses that darken at an adjustable frequency ranging from 70Hz to 90Hz in steps of 1Hz. We set the two devices to a common frequency of 80Hz. At this frequency, the flickering is imperceptible on paper, but create interference patterns with other frequencies, which makes them incompatible with reading on a computer screen or in a room lit by neon lights. All tests were therefore ran on paper, in a sufficiently bright room that did not require artificial lighting. Each student was tested individually during 5 sessions of 20 minutes spread over 5 consecutive weeks and comprising the following 5 conditions in pseudo-random order: glasses on, flickering at 80Hz; glasses off; lamp on, flickering at 80Hz; lamp on but not flickering; and natural light alone. This design made it possible to evaluate the placebo effect linked to the presence of the device alone with the impact of the device itself, as well as to see if one of the two device is more effective than the other.

The tests were conducted in a single-blind mode. The student did not know whether the lamp or the glasses were flickering or not. No comments regarding the functioning of the object or its facilitative potential were made by the experimenter. At the end of each session, the student was asked, among other things, if he or she had observed any flickering. The analysis of these responses indicates that the participants did not notice any difference between device on and device off conditions,  $\chi^2(1)=3.50$ ,  $p=0.062$ .

### ***Description of the tests***

Each session comprised the same 3 tests in random order.

### *Letter naming*

All 26 letters of the alphabet were presented twice and randomly distributed in 6 lines of 8 letters and a final line of 4 letters. Letters were printed in lower case, Calibri font, 14 point size, with 11 spaces between each letter and 1.5 line spacing on half an A4 page. Students were asked to name them aloud. Their overall reading time was measured using a stopwatch, switched on when the list of letters was presented to the student and stopped once the last letter had been spoken. We also reported errors.

### *Reading aloud a list of words*

Students were asked to read aloud a list of words, presented in columns, in lower case, Calibri font, 14 point size, 1.5 line spacing, on A4 paper. The manufacturers rely on the study by Le Floch and Ropars, which states that high-frequency flickering drastically reduces mirror confusions for letters b, d, p and q (Le Floch & Ropars, 2017). To test this hypothesis, we developed a list of 144 words, one third of which could be misread due to a confusion of those letters. We search the French lexicon for words forming a mirror pair, i.e. words in which if the substitution of a mirror letter (b, d p or q) by another mirror letter yields another word (e.g. bague – dague [ring - dagger]). We identified a list of 100 mirror pairs. Since the letters b, d, p and q are both visually and phonologically close, we also included phonological control pairs and visual control pairs. Because these letters are plosives, we selected the letters t, c (when pronounced /k/) and g (when pronounced /g/), which are also plosives and phonologically close, but bear little visual similarity to each other. By searching the lexicon, we obtained a list of 62 phonological word pairs, such that substitution of one such letter by another resulted in another word (e.g. grue – crue [crane – raw]). For the visually similar pairs, we used the similarity matrix obtained by Agrawal et al (Agrawal et al., 2020) to select the following pairs of similar letters f/l, r/v and n/h. Using a procedure similar to the one described above, we obtained 66 visual word pairs (e.g. localiser – focaliser [localise – focus]).

From these three lists, we selected 24 word pairs in each, so that the words in the three lists were matched according to their length, frequency, bigram frequency, number of neighbours (calculated

using OLD20), position of the substituted letter (first letter of the word or middle of the word), number of syllables, number of phonemes, orthographic CV structure and phonological CV structure. Those 144 words were presented in random order.

#### *Reading aloud a short text*

To measure students' reading fluency, we asked them to read a text called "Marianne" and comprising 295 words spread over 4 paragraphs whose sentences were all syntactically correct. Designed as a screener for various subtypes of dyslexia, it contained regular words, irregular words and pseudo words. Students were given 5 minutes to read as much of this text as possible. The same text was used for all sessions. We measured reading time when it was below 5 minutes, and the number of errors, thus allowing us to compute fluency as the number of correctly read items per minute.

#### **Data analysis**

All sessions were recorded to allow the tests to be rated by an external observer blind to the reading condition, thus ensuring that the rating was neutral. Two independent observers scored all the sessions and their results were more than 95% consistent. For the frequentist analysis, we used the following mixed effects models for each exercise:

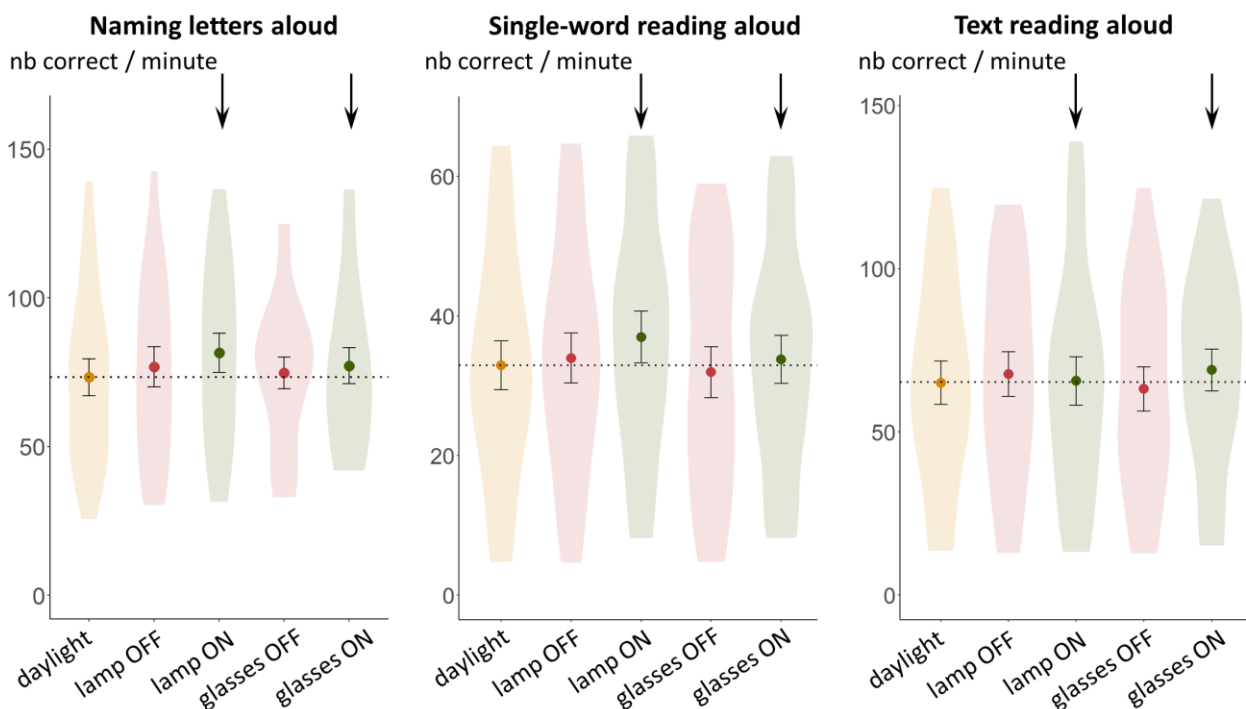
$$nb\_of\_correct\_responses\_per\_minute \sim order + condition + (1|participant)$$

The test order covariate (1-5) was added to the model to capture a putative learning effect, as the same tests were repeatedly used. Condition was a 5-level factor reflecting the condition in which the session took place (glasses or lamp flickering, glasses or lamp not flickering or natural light).

Bayesian analysis was carried out to assess the evidence for or against the hypothesis that the device on/off status had no effect on performance. We ran exactly the same model as the one used in the frequentist analysis.

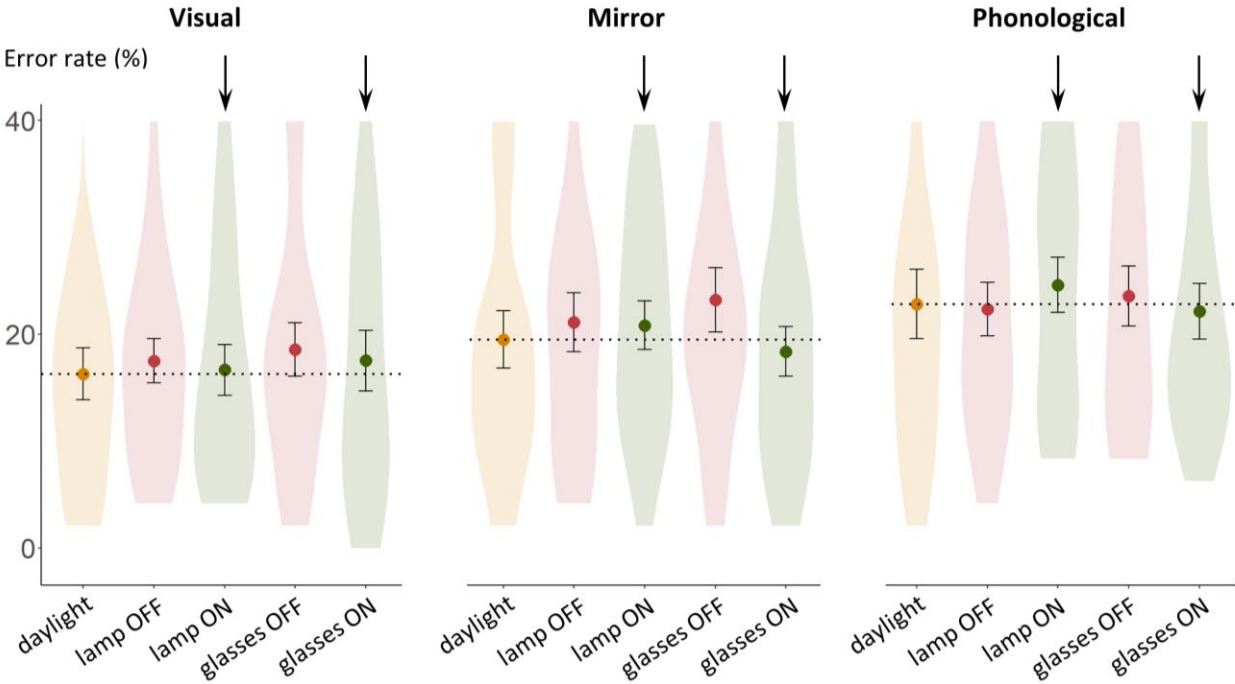
## Results

The distributions, across participants, of the number of correct answers per minute as a function of test condition and their means for each exercise are presented in Figure 4. Results were identical for both mixed-effects models. We found a positive effect of test order, with performance improving over time and repeated testing: letter naming:  $F(1,83.00)=5.55$ ,  $p=0.021$ ,  $BF=2.9$ ; list of words:  $F(1,81.99)=35.46$ ,  $p<0.01$ ,  $BF>100$ ; short text:  $F(1,83.00)=56.67$ ,  $p<0.001$ ,  $BF>100$ . The main effect of condition was never significant (letter naming:  $F(4,83.00)=1.21$ ,  $p=0.31$ ,  $BF=0.18$ ; list of words:  $F(1,81.99)=0.85$ ,  $p=0.50$ ,  $BF=0.11$ ; short text:  $F(4,83.00)=1.23$ ,  $p=0.303$ ,  $BF=0.18$ ). Thus, lighting conditions did not significantly influence the results obtained by students. This conclusion was supported by Bayes factors smaller than 1/3, corresponding to substantial evidence in favour of the null hypothesis.



**Figure 4.** Lack of effect of high-frequency flickering (80Hz) using either a lamp or glasses on 22 dyslexic children. Each point represents the mean score for each condition, and error bars represent  $\pm 1$  standard error of the mean. For reference, the dotted line shows the mean of the daylight condition and two arrows indicate the conditions under which the lamp or glasses were lit.

We next focused on mirror errors in the word lists, since high-frequency flickering has been claimed to reduce the mirror confusions made by patients with dyslexia (Le Floch & Ropars, 2017). Given our design, we compared the error rate in the three word-pair categories (mirror, phonological, visual). We submitted them to a mixed-effect model using order as a covariable and condition (5 levels) and category (3 levels, visual, mirror or phonological) as factors. It confirmed the significant main effect of order,  $F(1,290.01)=12.35$ ,  $p<0.001$ ,  $BF=47$ , as well as significant differences between categories of words  $F(2,290.01)=28.49$ ,  $p<0.001$ ,  $BF>100$ . As shown in Figure 5, the error rate for mirror-confusable words (average = 20.5%) fell in between the error rate for visually confusable words (17.3%) and for phonologically confusable words, which was the highest (23.0%). This result suggests that phonological similarity, more than left-right inversion or visual confusion, was the main source of errors for our participants. Crucially, however, there was again no main effect of lighting conditions  $F(4,290.04)=1.94$ ,  $p=0.10$ ,  $BF=0.17$ , and no significant interaction between condition and category  $F(8,290.01)=0.75$ ,  $p=0.65$ ,  $BF=0.028$ : lamps and glasses, whether on or off, had no effect on reading fluency for different types of error-inducing words.



**Figure 5.** Lack of effect of high frequency flickering (80Hz) using either a lamp or glasses on mirror errors on single -words reading. The graphs show the mean, standard error, and distribution of error rates across 22 dyslexic participants, separately for words that could be confused with another word by visual confusions (left), mirroring of a letter b d p q (middle), or phonological confusions (right). For reference, the dotted line shows the mean of the daylight condition and two arrows indicate the conditions under which the lamp or glasses were lit.

In summary, at the group level, we observed no significant impact of high-frequency flickering on letter, word, or text reading fluency. Individual analysis revealed no consistent improvement with the glasses or lamp. However, it remains possible that those devices may be helpful in a small number of specific cases, similar to those of McCloskey (2000), Pflugshaupt (2007) and Vannuscorps (2022). While this hypothesis is difficult to evaluate without testing an extremely large population, in experiment 4 we endeavoured to identify dyslexia cases who claimed to be helped by those devices – and then rigorously test if the effect was real or a placebo.

## **EXPERIMENT 4 - SINGLE CASE STUDY**

### **Method**

#### ***Participants***

A call for volunteers on social networks identified two participants, FAP and CT. Both were more than two standard deviations away from one of the two speed or accuracy variables in the Alouette test. For FAP, we used the adult norms, described in the article by Cavalli et al. (2018).

#### ***FAP***

FAP was 27 y.o. self-employed scientific illustrator. Reading has been difficult for her since the beginning of primary school due to dyslexia. She received speech therapy from ages 6 to 18. She has no other diagnosed learning disability.

FAP's reading experience is hindered by a hazy and glittering vision, leading to difficulty identifying words, which makes her very tired. However, wearing the glasses for over a year has improved her daily comfort, enabling longer reading sessions, and enhancing comprehension.

CT

CT was a 12 y.o. girl. Diagnosed as dyslexic at the end of 3<sup>rd</sup> grade, she is followed by a speech therapist for weekly sessions. She has been wearing the above glasses for two years now and reports a clear improvement in her reading skills. Without the glasses, she reports that the words stick together and overlap, while this is no longer the case when she wears them.

### ***Procedure***

We used a 2 x 2 design to contrast the genuine effect of glasses (on/off) with the placebo effect of believing that the glasses were on/off. The former variable is hereafter referred to as the objective variable: on different sessions, the glasses were either on and set to the participant's self-selected favourite frequency, or off. For the subjective variable, the different sessions were introduced with sentences such as "the glasses are now set to the frequency you usually use, so they should help you" versus "the glasses are now set to a different frequency than that you usually use, so they should not help you". These two objective/subjective factors were crossed in a 2x2 factorial design, repeated twice, for a total of 8 sessions. The order of the sessions was reversed between FAP and CT. A training session was also conducted to familiarise the participants with the instructions and the different tests.

During each session, participants were asked to perform three different tests, all carried out on paper. In order to collect reading and decision times, the whole session was filmed with a 360p, 16:9, 30 fps camera.

### ***Single words reading aloud***

To test single-word reading, we used a sub-list from experiment 3. We classified words used in experiment 3 by category (mirror, visual or phonological pairs), length (short: 4-5 letters / medium: 6 letters / long: 7-8 letters) and trap position (beginning of the word or middle of the word). Within this classification, we retained two pairs of words, resulting in a list of 36 word pairs. In each session, half of this list (36 words) was presented in random order. Overall, since each objective x subjective



condition was presented twice to each participant, the entire list was read in each condition of the 2x2 design.

To determine the reading time of each item, the list was presented as follows: the 36 words were divided into 4 lines of 9 words, written in Calibri 14 on a blank sheet of paper. A mask allowed one word to appear after another, hiding the rest of the list. The reading time was determined as the difference between the time when the word appeared entirely in the cache (determined manually using frame-by-frame video analysis) and the time when the participant started to say it aloud (determined manually using the audio analysis software Audacity).

#### *Text reading aloud*

The design of the present experiment did not allow for the reuse of the same text (“Marianne”) as in experiment 3: as all eight sessions were carried out within a two-hour interval, the repetition effect would have been massive. Instead, the texts used here were taken from the ALECTOR corpus (Gala et al., 2020), which lists reading resources for children from 2<sup>nd</sup> to 4<sup>th</sup> grade. We extracted all texts suitable for children in 4<sup>th</sup> grade, an age reasonably younger than that of our participants, and cut them up into slices of about 200 words, resulting in a corpus of 28 texts (13 extracts from novels and stories and 15 extracts from science documentaries).

During each session, participants were asked to read 3 texts (4 texts were used for the training session), written in Calibri 14 with 1.5 line spacing. For each text, participants were asked to go as far as possible in one minute. We recorded reading fluency (number of words correctly read in one minute) and error rate.

#### *Sentence comprehension*

The sentence comprehension test was adapted from the Score Aphasiologique de la Salpêtrière (SAS) listening comprehension test. Originally, it is a listening comprehension test in which the experimenter reads aloud a sentence and the participant has to choose the relevant image among 4. The original

test is composed of 90 items divided into 7 categories: active sentences with one or two distractors on the picture ("The policeman pursues the thief"), passive sentences with one or two distractors on the picture ("The thief is pursued by the policeman"), positioning of geometric shapes in relation to each other ("The rectangle is to the left of the square"), subject relatives with the pronoun "qui" in French ("The truck which is following the car is black") and semantically reversible object relatives with the pronoun "que" ("The truck that the car is following is black"). For the purposes of our experiment, we added 38 items to obtain a total of 14 sentences per session, 2 sentences from each category. The original test had 18 geometric shape positioning items, so 2 of these were removed.

To turn this test into a reading comprehension test, participants were asked to read each sentence (aloud or silently) and then point to the appropriate picture. The test was administered using a binder in which the pages alternated between sentences and associated images. Thus, participants no longer had the sentence in front of them when they pointed to the image. The decision time was measured as the difference between the moment when the participant could see all four images and the moment when her finger touched one of them.

### ***Data analysis***

To analyse these data, we run ANCOVAs for each participant and each exercise using a 2x2 factorial design with objective (what we did) and subjective (what we said we did) binary variables:

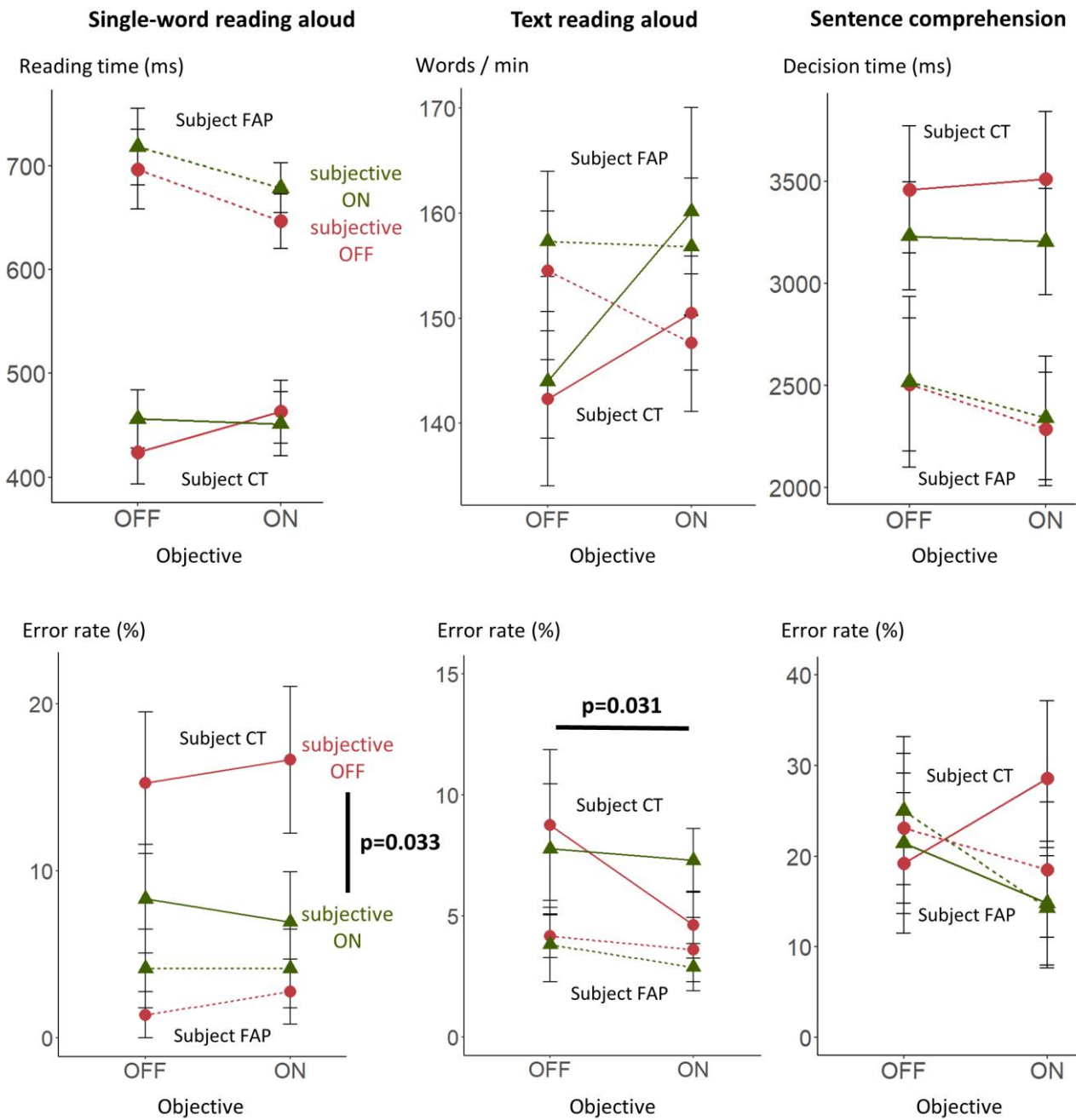
$$dv \sim order + subjective * objective$$

The dependent variable was fluency or error rate for text reading, reading time or error rate for words reading, and decision time or error rate for sentence comprehension. We added the covariate of test order (1-8) to capture a putative learning effect. We excluded trials with reading times more than three standard deviations away from the mean for each participant (less than 3% for each participant in each test). For word reading and sentence comprehension, we only considered reading time on correct trials. The exact same model was used for Bayesian analysis.

## Results

### *Single-word reading aloud*

Overall, CT was faster than FAP (average reading time of 447 ms for CT and 688 ms for FAP) but she made more errors (error rate of 12.0% for CT and 3.19% for FAP). Figure 6 shows reading times and error rates in each condition of the experimental design. On reading times, we found a significant effect of temporal order in both participants (FAP:  $F(1,268)=10.66$ ,  $p=0.001$ ,  $BF=24$ ; CT:  $F(1,245)=12.14$ ,  $p<0.001$ ,  $BF=42$ ), but no significant effect of either the subjective (FAP:  $F(1,268)=0.77$ ,  $p=0.38$ ,  $BF=0.20$ ; CT:  $F(1,245)=0.19$ ,  $p=0.67$ ,  $BF=0.15$ ) or the objective variable (FAP:  $F(1,268)=0.008$ ,  $p=0.93$ ,  $BF=0.16$ ; CT:  $F(1,245)=3.93$ ,  $p=0.05$ ,  $BF=0.89$ ). Their interaction was also not significant (FAP:  $F(1,268)=0.049$ ,  $p=0.83$ ,  $BF=0.19$ ; CT:  $F(1,245)=0.49$ ,  $p=0.48$ ,  $BF=0.24$ ). Thus, reading times, for both participants, did not differ across conditions.



**Figure 6.** Lack of effect of high frequency flickering (FAP: 87Hz or CT: 82Hz) in two single participants claiming to be helped by flickering glasses. Scores in the different exercises for both participants. Significant differences have been highlighted with *p* value.

On accuracy, we found no significant effect of temporal order (FAP:  $F(1,277)=1.23$ ,  $p=0.27$ ,  $BF=0.24$ ; CT:  $F(1,279)=0.21$ ,  $p=0.65$ ,  $BF=0.16$ ) or of the objective variable (FAP:  $F(1,277)=0.58$ ,  $p=0.45$ ,  $BF=0.18$ ; CT:  $F(1,279)=0.045$ ,  $p=0.83$ ,  $BF=0.15$ ). But we found a small effect of the subjective variable on CT's error rate (FAP:  $F(1,277)=1.03$ ,  $p=0.31$ ,  $BF=0.21$ ; CT:  $F(1,279)=4.60$ ,  $p=0.033$ ,  $BF=1.2$ ). CT made fewer

errors in single-word reading when told that the glasses were on. This effect was quite modest as the subjective variable only explained 1,6% of the variance in error rates and the Bayes factor was close to one. The lack of interaction with the objective variable (FAP:  $F(1,277)=0.11$ ,  $p=0.74$ ,  $BF=0.19$ ; CT:  $F(1,279)=0.17$ ,  $p=0.68$ ,  $BF=0.18$ ) suggested that this effect was unchanged whether the glasses were actually on or off – a pure placebo effect.

### ***Text reading aloud***

Overall, FAP read on average 154 words correctly in 1 minute, and CT 151. FAP's accuracy was slightly higher than CT's as she made 3.8% errors compared to 7.3% for CT. The results of the ANCOVA highlighted the lack of effect of glasses on reading speed. None of the variables in the model reached significance for both participants and all Bayes factors were smaller than one. In contrast, on error rates, there was a small but significant interaction between objective and subjective variables, in CT only (FAP:  $F(1,19)=0.078$ ,  $p=0.78$ ,  $BF=0.48$ ; CT:  $F(1,19)=5.20$ ,  $p=0.034$ ,  $BF=2.5$ ). In a simple effect analysis, a small effect of the objective variable (a reduction of error rates when the glasses were on rather than off) was found only when CT was told that the glasses are off,  $F(1,9)=6.56$ ,  $p=0.031$ ,  $BF=3.8$ . While this effect goes in the correct direction, it should be noted that it would not resist a correction for multiple comparisons and, most importantly, if it was a genuine effect rather than a false positive, it is hard to see why it would not be replicated in the “subjective on, objective on” condition, which yielded more errors.

### ***Sentence comprehension***

Finally, we were interested in the impact of glasses on reading comprehension. Overall, FAP responded faster than CT (FAP: 2.6s; CT: 3.7s), but both made the same amount of errors (FAP: 20.2%; CT: 21.1%). None of the effects in the ANCOVA on decision time reached significance for either participant, whether this was the effect of the subjective condition (FAP:  $F(1,82)=7.00 \cdot 10^{-3}$ ,  $p=0.93$ ,  $BF=0.23$ ; CT:  $F(1,81)=0.93$ ,  $p=0.34$ ,  $BF=0.33$ ), the effect of the objective condition (FAP:  $F(1,82)=0.075$ ,  $p=0.79$ ,  $BF=0.26$ ; CT:  $F(1,81)=0.31$ ,  $p=0.58$ ,  $BF=0.27$ ) or their interaction (FAP:  $F(1,82)=8.2 \cdot 10^{-4}$ ,  $p=0.98$ ,

BF=0.30; CT:  $F(1,81)=4.00 \cdot 10^{-3}$ ,  $p=0.95$ , BF=0.30). This result was strengthened by Bayesian analysis in both participant as for both decision time all Bayes factors were smaller than 1/3. Similar results were observed for error rates.

In summary, two patients claiming benefits from flickering glasses showed minimal or no objective effect. In FAP, no effect of objective or subjective variables was found in both tests. CT was influenced by a placebo effect during single-word reading and, while a small improvement of sentence reading accuracy was found when the glasses were on, the fact that it was small, only appeared as an interaction with the subjective variable, and only in a single test, suggests that it was likely a false positive.

## **GENERAL DISCUSSION**

Our aim, through these different experiments, was to assess the impact of low and high frequency flickering on reading. We found no major effect of low-frequency flickering, either in adults or in dyslexic children: periodically refreshing bottom-up inputs did not facilitate reading. We only found that low-frequency flicker slightly biased adults towards pseudowords in the lexical decision task. These results confirm that reading difficulty profiles such as those described by McCloskey and Pflugshaupt (McCloskey & Rapp, 2000a; Pflugshaupt et al., 2007), who were helped by low-frequency flickering, are quite rare.

Regarding high-frequency flickering, our data favours the absence of effect of the lamp or the glasses, the performance of the students being very similar whatever the reading test we proposed. Our findings, of course, should not be taken to imply that those devices may never be helpful to some readers. However, they stand in stark contrast with marketing claims that they facilitate reading for 90% of dyslexic children (Atol les Opticiens, communication personnelle, 16 mars 2022). We also found no impact of both devices on the rate of mirror confusions of students, a result that contrasts with

prior suggestions on the usefulness of high-frequency flickering to reduce mirror image formation in dyslexics (Le Floch & Ropars, 2017).

Even in two dyslexic participants who felt helped by the glasses, we found no major improvement in either fluency or comprehension. We only observed a weak placebo effect in our youngest participant, whose accuracy on word reading improved when she was told that the glasses were on. She also showed a slight improvement in text reading accuracy when the glasses were actually on but this effect was only in one test and, inexplicably, only when she was told that the glasses were off.

While these findings contrast with those described by Le Floch and Ropars, they are consistent with the literature on flicker contrast sensitivity, which shows that dyslexics are no more sensitive than normal readers to low and high frequency flickering (Cornelissen et al., 1995; Williams et al., 2003). These studies, however, only focused on low-level visual perception and did not investigate the impact on reading. To our knowledge, the present study is the first to assess the effects of high and low frequency flickering on reading in normal readers and in a group of dyslexic children. While revising the present paper, we became aware of a preprint that draws similar conclusions (Lapeyre et al., 2023). Lapeyre et al. investigated the impact of the third flickering device currently on the market, the Lili lamp, in dyslexic adults and age-matched controls. They first assessed reading deficits using standardised tests, as well as measured visual acuity and ocular dominance. These measures already call into question Le Floch and Ropars' hypothesis that dyslexia is linked to a dominant-eye deficit, since 87% of their dyslexic participants showed normal ocular dominance. Furthermore, flickering light had no significant impact on subsequent tests of sentence reading speed and text reading comprehension.

While further studies could possibly identify a subtype of dyslexia that would be sensitive to flicker, the weight of the evidence, across 4 successive experiments, indicates that flicker is not a viable solution to the reading difficulties of most, if not all, individuals. We find this conclusion unsurprising for two reasons: first, the slim evidence previously presented in support of the efficacy of flickering (Le

Floch & Ropars, 2017); and second, the overwhelming brain-imaging evidence that reading acquisition and reading deficits occur in the cortex rather than the retina, and involve a broad hierarchy of areas, most of which lie above the level of invariance where flicker would be expected to have an effect (Dehaene et al., 2015; Feng et al., 2020b). Nevertheless, the present research highlights the importance and the feasibility of using the cognitive psychology of reading to evaluate the claims of device manufacturers in this field. Indeed, it is hard to understand why the burden of proof does not lie with the manufacturers themselves, prior to selling their products, as in the medical domain. We hope that the present work may constitute a small step in making evidence-based psychology the future norm.



## CONCLUSION AND FUTURE RESEARCH DIRECTIONS

## **GENERAL CONCLUSION**

For this doctoral research, our primary objectives were twofold: to enhance dyslexia screening methods in France by crafting diagnostic tools and to assess the efficacy of existing remediation tools available in the market. We envisioned that the outcomes of our work could be readily implemented through our collaborative partnership with the CERENE schools and diagnostic center. In this summary, I will outline our key contributions to the fields of reading and dyslexia sciences and revisit the project's deliverables.

### **Contributions to the science of reading**

#### ***How does variability in typical reading affected lexical decision results in students with the same amount of education?***

In my introduction, I mentioned our theoretical aim of studying how variability in typical reading affected lexical decision results in students with the same amount of education. To this end, I developed a lexical decision task, which was administered to 1,501 6<sup>th</sup> graders representative of the French population as part of the National Assessments organised each year by the Ministry of Education. Observed through the prism of the level of reading fluency measured during these same assessments, my results enabled me to confirm the variation in reading fluency at this age, which ranged from adult-like to lagging behind 3rd grade expectations (Andreu et al., 2022). Furthermore, this study replicated some of the hallmark phenomena in visual word recognition, namely: lexicality effect (B. S. Weekes, 1997; Seidenberg et al., 1984; Forster & Chambers, 1973; Juphard et al., 2004a), length effect (Bijeljac-Babic et al., 2004a; Ferrand et al., 2010; New et al., 2006a; B. S. Weekes, 1997) and frequency effect (Andrews, 1992, 1989; Forster & Chambers, 1973; McCann & Besner, 1987; Seidenberg et al., 1984; Burani et al., 2002). Their interaction with fluency level revealed profiles of readers similar to those observed in many cross-sectional studies, with the best readers being faster and more accurate at making lexical decisions, and their effects of length, lexicality and frequency being less marked than in the least fluent students, who showed a greater effect of length, even for

high frequency words (Bijeljac-Babic et al., 2004a; Burani et al., 2002; Schmalz et al., 2013a; Schröter & Schroeder, 2017; Zoccolotti et al., 2005a, 2009b). While we were unable to formally exclude dyslexic students from our participant pool or conduct multiple test sessions to assess test-retest reliability, the substantial size of our sample and the inclusion of students spanning a broad spectrum of reading abilities render our study a valuable addition to the literature on reading across diverse orthographies.

### ***What kind of errors are typically made by normal readers?***

During this PhD, I was interested in the different types of reading errors made by typical readers. Using again the lexical decision, I first looked at how different types of pseudo-words were processed by the 1501 6<sup>th</sup> graders described above. In general agreement with previous research, I found that stimuli that were most similar to real French words yielded that highest error rate (Bergmann & Wimmer, 2008; Grainger et al., 2012). My results also showed that almost no students, even the most fluent, perfectly master French contextual rules, and that contrary to what we were expecting, mirror letter substitutions are as hard to process as single letter substitutions.

In a second study, I visited 812 pupils from 1<sup>st</sup> to 5<sup>th</sup> grade to have them read the text of the Mariette, a read-aloud exercise designed to screen for different subtypes of dyslexia. Generally speaking, I observed a significant drop in the total percentage of errors and an increase in reading speed compatible with the results of the literature (Al Otaiba et al., 2009; Biemiller, 1977; Lervag & Hulme, 2009; Nese et al., 2013). By characterising each of the errors individually, I then looked at the development of different types of error throughout primary school. I found that a substantial portion of visual errors (including omissions, substitutions, and letter additions), and certain surface errors such as the pronunciation of silent final letters or mispronunciation of specific words tend to diminish significantly as children acquire reading skills. However, other error types such as regularizations, functional word confusions, and the application of contextual rules persist even in 5th graders. It's noteworthy that this categorization of errors has not been comprehensively explored in a large sample of typical young readers until now.

This novel categorization of errors holds significant implications for the field of applied research in French literacy. It not only provides a clear understanding of the normal progression of reading errors at each grade level but also informs the development of targeted instructional methods. Specifically, it sheds light on the aspects of reading, such as the mastery of conceptual rules, that still require attention and refinement in pedagogical approaches, contributing to the science of reading as well as dyslexia.

### **Contributions to the science of dyslexia**

The applied aspect of this thesis also provided an opportunity to enhance existing screening tools for dyslexia in French, in particular through the development of a screening tool for different subtypes of dyslexia. It was in this context that we developed the Mariette, a 294-word non-meaning text containing irregular words, pseudowords and transposable words. This text, designed as an initial screener, enables us to situate the pupils' various errors among those of their peers, identify any difficulties they may have and adapt the rest of the assessment accordingly to confirm the diagnosis. This is the first time that a test of this type, offering a precise characterization of children's specific errors, has been produced in French for primary school pupils. With this test, we bring to research into dyslexia in French the possibility of better characterizing the profiles of dyslexic participants, and to international studies a method for developing such tools and quantifying pupils' errors.

Another contribution of my work to the science of dyslexia concerns our finding that high or low frequency flickering is useless in facilitating reading in most dyslexics. This result, which we obtained from 375 normal adult readers and a cohort of 20 dyslexic pupils, calls into question a recent idea according to which dyslexia could be corrected using flickering devices (Le Floch & Ropars, 2017) and on which French manufacturers had relied to design lamps and glasses supposed to make reading easier for dyslexic children. Similar conclusions were drawn by Lapeyre and colleagues in a recent study conducted in adults (Lapeyre et al., 2023). These results not only add to the scientific literature on

dyslexia, but also constitute the first solid scientific studies on the subject. They are even more important as they will enable speech therapists and parents to make more informed choices in the support of dyslexic children.

## **Our deliverables**

All of this work was carried out with a view to providing CERENE and, more broadly, practitioners and teachers with applicable solutions for screening and characterising reading difficulties. Two tools were designed to this aim. The lexical decision has been standardised for 6<sup>th</sup> graders and can now be accessed free of charge online, on computers and mobile phones. At the end of each test, which lasts around 3 minutes, automatic feedback is generated, showing where the pupil's strengths and weaknesses lie. Regular use of the test allows you to see how the student is developing and to measure their progress. The Mariette has been standardised for pupils from 1<sup>st</sup> to 5<sup>th</sup> grade and will soon be available to teachers and practitioners who wish to use it. A user guide including norms and rules to make it easier to mark errors, will accompany the test.

Finally, with our study on flickering, we have provided CERENE's paramedical professionals with arguments in response to parents' questions about reading tools that are currently on the market with the unfounded promise of helping almost 90% of dyslexic children.

## **FUTURE RESEARCH DIRECTIONS**

I will conclude by mentioning my prospects for future research in applied science for dyslexia diagnosis and remediation.

### **Continuing to develop a battery of tests focusing on children's specific errors**

The development of the Mariette is part of a wider project drive by CERENE and our research team to develop a dyslexia screening battery that focuses on children's precise errors. Prior to this thesis, this battery already included a word identification test, the Malabi (Potier Watkins et al., 2023), standardised for French 6<sup>th</sup> and 7<sup>th</sup> graders. A first limitation of this work concerns the range of levels

for which these tests have been standardised. Indeed, it would be interesting to administer the Malabi to pupils for whom the Mariette results seem to indicate a specific disorder such as letter position dyslexia or attentional dyslexia. However, analysis of the results will not allow any conclusions to be drawn, as there is no Malabi norm for primary school pupils. A first objective in continuing this work could be to standardise these two tests from 1<sup>st</sup> to 12<sup>th</sup> grade. Having the same pupils take both tests would also make it possible to measure quantitatively the reliability of the clues given by the Mariette. The complexity of the scoring is also a limiting factor for the inclusion of such tests in current speech and language therapy assessments. It would therefore be interesting to develop a speech recognition system which, linked to the database of errors that we have already established, would enable virtually automatic scoring. Such a tool would certainly enable our battery to be disseminated more widely.

### **Large scale studies to measure the prevalence of each of these types of dyslexia in French and in other languages**

As mentioned in the introduction, Naama Friedmann's work has highlighted different subtypes of dyslexia in a vast variety of languages (Friedmann, Kerbel, et al., 2010; Friedmann & Haddad-Hanna, 2012, 2014; Friedmann & Rahamim, 2007a; Güven & Friedmann, 2019; Kohnen et al., 2012; Potier Watkins et al., 2023; Rayner et al., 1989). However, there are currently no studies measuring the prevalence of these different disorders in the general population. So once our battery has been standardised for pupils of the same age, it would be interesting to conduct large-scale studies in French, and also in the different languages for which these different subtypes of dyslexia have been identified. This could give rise to comparisons between countries, to understand how the opacity of the language influences the representation of these different subtypes of dyslexia.

### **Find personalised solutions based on the dyslexia profile**

Another direction in which we could pursue our research concerns remedial solutions for dyslexia, and in particular, the development of specific solutions for each subtype of dyslexia. Previous research have shown that the use of an attentional window drastically reduced errors in attentional dyslexics

(Friedmann, Kerbel, et al., 2010), while following words with the finger greatly helped patients with letter-position dyslexia (Friedmann & Rahamim, 2014). Maybe flickering devices could also help some children with a specific profile. Once enough cases of each subtype of dyslexia have been identified, new remediation practices need to be designed and tested. From an applied science perspective, disseminating successful methods to practitioners, teachers and parents should also be one of our objectives, to improve the care of dyslexic children.

## REFERENCES

- Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision : Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology*, *99*(2), 245-264. <https://doi.org/10.1348/000712607X224478>
- Agrawal, A., Hari, K., & Arun, S. (2020). A compositional neural code in high-level visual cortex can explain jumbled word reading. *eLife*, *9*, e54846. <https://doi.org/10.7554/eLife.54846>
- Ahr, E., Houdé, O., & Borst, G. (2016). Inhibition of the mirror generalization process in reading in school-aged children. *Journal of Experimental Child Psychology*, *145*, 157-165. <https://doi.org/10.1016/j.jecp.2015.12.009>
- Al Otaiba, S., Petscher, Y., Pappamihiel, N. E., Williams, R. S., Dyrland, A. K., & Connor, C. (2009). Modeling oral reading fluency development in Latino students : A longitudinal study across second and third grade. *Journal of Educational Psychology*, *101*(2), 315-329. <https://doi.org/10.1037/a0014698>
- Alexander, A. W., Andersen, H. G., Heilman, P. C., Voeller, K. K. S., & Torgesen, J. K. (1991). Phonological awareness training and remediation of analytic decoding deficits in a group of severe dyslexics. *Annals of Dyslexia*, *41*(1), 193-206. <https://doi.org/10.1007/BF02648086>
- Altarelli, I., Leroy, F., Monzalvo, K., Fluss, J., Billard, C., Dehaene-Lambertz, G., Galaburda, A. M., & Ramus, F. (2014). Planum temporale asymmetry in developmental dyslexia : Revisiting an old question: Planum Temporale Asymmetry in Dyslexia. *Human Brain Mapping*, *35*(12), 5717-5735. <https://doi.org/10.1002/hbm.22579>
- Andreu, S., Ali, L., Bret, A., Santos, R., Monestrol, H. D. de, Lambert, K., M'Bafoumou, A., Paillet, V., Rocher, T., & Vourc', R. (2021). 800 000 élèves évalués en début de sixième en 2020. 1. <https://doi.org/10.48464/ni-21-03>
- Andreu, S., Ballereau, M.-A., Ali, L., Bret, A., Santos, R., Monestrol, H. D. de, Heidmann, L., Lambert, K., M'Bafoumou, A., Paillet, V., Paul, A., Persem, É., Rocher, T., Rue, G., Stachowiak, J.-F., Salles, F., & Vourc', R. (2022). Évaluations de début de sixième en 2021. <https://doi.org/10.48464/ni-22-04>
- Andreu, S., Conceicao, P., Eteve, Y., & Vourc'h, R. (2023). Alors que l'année de CP permet de réduire les écarts de performances entre secteurs de scolarisation, les vacances scolaires les accentuent. *Note d'Information*, n° 23.17, DEPP.
- Andrews, S. (1989). Frequency and neighborhood effects on lexical access : Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 802-814. <https://doi.org/10.1037/0278-7393.15.5.802>
- Andrews, S. (1992). Frequency and neighborhood effects on lexical access : Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 234-254. <https://doi.org/10.1037/0278-7393.18.2.234>
- Araújo, S., Faisca, L., Bramão, I., Petersson, K. M., & Reis, A. (2014a). Lexical and Phonological Processes in Dyslexic Readers : Evidence from a Visual Lexical Decision Task: Lexical and Phonological Processes in Dyslexia. *Dyslexia*, *20*(1), 38-53. <https://doi.org/10.1002/dys.1461>
- Araújo, S., Faisca, L., Bramão, I., Petersson, K. M., & Reis, A. (2014b). Lexical and Phonological Processes in Dyslexic Readers : Evidence from a Visual Lexical Decision Task: Lexical and Phonological Processes in Dyslexia. *Dyslexia*, *20*(1), 38-53. <https://doi.org/10.1002/dys.1461>



- Arnold, D. H., Williams, J. D., Phipps, N. E., & Goodale, M. A. (2016). Sharpening vision by adapting to flicker. *Proceedings of the National Academy of Sciences*, *113*(44), 12556-12561. <https://doi.org/10.1073/pnas.1609330113>
- Atol les Opticiens. (2022, mars 16). *Atol les Opticiens continue d'innover sur le marché de l'optique et présente les lunettes électroniques intelligentes LEXILENS® destinées aux enfants et aux adultes dyslexiques* [Press release].
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *55*(2), 290-313. <https://doi.org/10.1016/j.jml.2006.03.008>
- Bachmann, C., & Mengheri, L. (2018). Dyslexia and Fonts : Is a Specific Font Useful? *Brain Sciences*, *8*(5), 89. <https://doi.org/10.3390/brainsci8050089>
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual Word Recognition of Single-Syllable Words. *Journal of Experimental Psychology: General*, *133*(2), 283-316. <https://doi.org/10.1037/0096-3445.133.2.283>
- Balota, D. A., Yap, M. J., & Cortese, M. J. (2006). Chapter 9 - Visual Word Recognition : The Journey from Features to Meaning (A Travel Update). In M. J. Traxler & M. A. Gernsbacher (Éds.), *Handbook of Psycholinguistics (Second Edition)* (p. 285-375). Academic Press. <https://doi.org/10.1016/B978-012369374-7/50010-9>
- Banfi, C., Kemény, F., Gangl, M., Schulte-Körne, G., Moll, K., & Landerl, K. (2018). Visual attention span performance in German-speaking children with differential reading and spelling profiles : No evidence of group differences. *PLOS ONE*, *13*(6), e0198903. <https://doi.org/10.1371/journal.pone.0198903>
- Barca, L., Burani, C., Di Filippo, G., & Zoccolotti, P. (2006). Italian developmental dyslexic and proficient readers : Where are the differences? *Brain and Language*, *98*(3), 347-351. <https://doi.org/10.1016/j.bandl.2006.05.001>
- Beauvois, M. F., & Derouesne, J. (1979). Phonological alexia : Three dissociations. *Journal of Neurology, Neurosurgery & Psychiatry*, *42*(12), 1115-1124. <https://doi.org/10.1136/jnnp.42.12.1115>
- Beech, J. R., & Awaida, M. (1992). Lexical and nonlexical routes : A comparison between normally achieving and poor readers. *Journal of learning disabilities*, *25*(3), 196-206.
- Bergmann, J., & Wimmer, H. (2008). A dual-route perspective on poor reading in a regular orthography : Evidence from phonological and orthographic lexical decisions. *Cognitive Neuropsychology*, *25*(5), 653-676. <https://doi.org/10.1080/02643290802221404>
- Biemiller, A. (1977). Relationships between Oral Reading Rates for Letters, Words, and Simple Text in the Development of Reading Achievement. *Reading Research Quarterly*, *13*(2), 223. <https://doi.org/10.2307/747307>
- Bijeljac-Babic, R., Millogo, V., Farioli, F., & Grainger, J. (2004a). A developmental investigation of word length effects in reading using a new on-line word identification paradigm. *Reading and Writing*, *17*(4), 411-431. <https://doi.org/10.1023/B:READ.0000032664.20755.af>

- Bijeljac-Babic, R., Millogo, V., Farioli, F., & Grainger, J. (2004b). A developmental investigation of word length effects in reading using a new on-line word identification paradigm. *Reading and Writing, 17*(4), 411-431. <https://doi.org/10.1023/B:READ.0000032664.20755.af>
- Boden, C., & Giaschi, D. (2007). M-stream deficits and reading-related visual processes in developmental dyslexia. *Psychological Bulletin, 133*(2), 346-366. <https://doi.org/10.1037/0033-2909.133.2.346>
- Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited : Universal structures plus writing system variation. *Human Brain Mapping, 25*(1), 92-104. <https://doi.org/10.1002/hbm.20124>
- Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia : The visual attention span deficit hypothesis. *Cognition, 104*(2), 198-230. <https://doi.org/10.1016/j.cognition.2006.05.009>
- Bosse, M.-L., & Valdois, S. (2003). Patterns of developmental dyslexia according to a multi-trace memory model of reading. *Current psychology letters, 10, Vol. 1, 2003*. <https://doi.org/10.4000/cpl.92>
- Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—A causal connection. *Nature, 301*(5899), 419-421. <https://doi.org/10.1038/301419a0>
- Braga, L. W., Amemiya, E., Tauil, A., Suguieda, D., Lacerda, C., Klein, E., Dehaene-Lambertz, G., & Dehaene, S. (2017). Tracking Adult Literacy Acquisition With Functional MRI : A Single-Case Study: Tracking Adult Literacy Acquisition With fMRI. *Mind, Brain, and Education, 11*(3), 121-132. <https://doi.org/10.1111/mbe.12143>
- Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., Lyytinen, H., Brandeis, D., & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences, 107*(17), 7939-7944. <https://doi.org/10.1073/pnas.0904402107>
- Brunsdon, R., Coltheart, M., & Nickels, L. (2006). Severe developmental letter-processing impairment : A treatment case study. *Cognitive Neuropsychology, 23*(6), 795-821. <https://doi.org/10.1080/02643290500310863>
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The Word Frequency Effect : A Review of Recent Developments and Implications for the Choice of Frequency Estimates in German. *Experimental Psychology, 58*(5), 412-424. <https://doi.org/10.1027/1618-3169/a000123>
- Burani, C., Marcolini, S., & Stella, G. (2002). How Early Does Morphological Reading Develop in Readers of a Shallow Orthography? *Brain and Language, 81*(1-3), 568-586. <https://doi.org/10.1006/brln.2001.2548>
- Castles, A. (2006). The dual route model and the developmental dyslexias. *London Review of Education*. <https://doi.org/10.1080/13603110600574454>
- Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*(2), 149-180. [https://doi.org/10.1016/0010-0277\(93\)90003-E](https://doi.org/10.1016/0010-0277(93)90003-E)
- Castles, A., & Friedmann, N. (2014). Developmental Dyslexia and the Phonological Deficit Hypothesis. *Mind & Language, 29*(3), 270-285. <https://doi.org/10.1111/mila.12050>
- Cavalli, E., Colé, P., Leloup, G., Poracchia-George, F., Sprenger-Charolles, L., & El Ahmadi, A. (2018). Screening for Dyslexia in French-Speaking University Students : An Evaluation of the Detection

- Accuracy of the *Alouette* Test. *Journal of Learning Disabilities*, 51(3), 268-282.  
<https://doi.org/10.1177/0022219417704637>
- Chabanon, L., & Rosenwald, F. (2020). Journée défense et citoyenneté 2019 : Plus d'un jeune français sur dix en difficultés de lecture. *Note d'information - DEPP -Ministère de l'Éducation Nationale et de la Jeunesse*.
- Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behavior*, 18(2), 225-241. [https://doi.org/10.1016/S0022-5371\(79\)90136-1](https://doi.org/10.1016/S0022-5371(79)90136-1)
- Cohen, L. (2003). Visual Word Recognition in the Left and Right Hemispheres : Anatomical and Functional Correlates of Peripheral Alexias. *Cerebral Cortex*, 13(12), 1313-1333.  
<https://doi.org/10.1093/cercor/bhg079>
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area. *Brain*, 123(2), 291-307. <https://doi.org/10.1093/brain/123.2.291>
- Collis, N. L., Kohonen, S., & Kinoshita, S. (2013). The role of visual spatial attention in adult developmental dyslexia. *Quarterly Journal of Experimental Psychology*, 66(2), 245-260.  
<https://doi.org/10.1080/17470218.2012.705305>
- Coltheart, M. (2007). Modeling Reading : The Dual-Route Approach. In M. J. Snowling & C. Hulme (Éds.), *The Science of Reading, A Handbook* (Blackwell publishing, p. 6-23).
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic, *Attention and performance* (p. 535-555).
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001a). DRC : A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256.  
<https://doi.org/10.1037/0033-295X.108.1.204>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001b). *DRC: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud*. 53.
- Cornelissen, P., Richardson, A., Mason, A., Fowler, S., & Stein, J. (1995). Contrast sensitivity and coherent motion detection measured at photopic luminance levels in dyslexics and controls. *Vision Research*, 35(10), 1483-1494. [https://doi.org/10.1016/0042-6989\(95\)98728-R](https://doi.org/10.1016/0042-6989(95)98728-R)
- Cornell, J. M. (1985). Spontaneous mirror-writing in children. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 39(1), 174-179. <https://doi.org/10.1037/h0080122>
- Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables : An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology*, 60(8), 1072-1082. <https://doi.org/10.1080/17470210701315467>
- Davies, R., Rodríguez-Ferreiro, J., Suárez, P., & Cuetos, F. (2013). Lexical and sub-lexical effects on accuracy, reaction time and response duration : Impaired and typical word and pseudoword reading in a transparent orthography. *Reading and Writing*, 26(5), 721-738. <https://doi.org/10.1007/s11145-012-9388-1>
- Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding : Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 535-557. <https://doi.org/10.1037/0096-1523.32.3.535>
- De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The Effect of Word Length and Other Sublexical, Lexical, and Semantic Variables on Developmental Reading Deficits. *Cognitive and Behavioral Neurology*, 21(4), 227-235. <https://doi.org/10.1097/WNN.0b013e318190d162>

- Dehaene, S. (2009). *Reading in the brain*. Penguin Viking.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15(6), 254-262. <https://doi.org/10.1016/j.tics.2011.04.003>
- Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate : Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews. Neuroscience*, 16(4), 234-244. <https://doi.org/10.1038/nrn3924>
- Dehaene, S., Le Clec'H, G., Poline, J.-B., Le Bihan, D., & Cohen, L. (2002). The visual word form area : A prelexical representation of visual words in the fusiform gyrus. *NeuroReport*, 13(3), 312-325.
- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4(7), 752-758.
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How Learning to Read Changes the Cortical Networks for Vision and Language. *Science*, 330(6009), 1359-1364. <https://doi.org/10.1126/science.1194140>
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., & Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences*, 103(38), 14240-14245. <https://doi.org/10.1073/pnas.0606302103>
- Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form : Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLOS Biology*, 16(3), e2004103. <https://doi.org/10.1371/journal.pbio.2004103>
- Déjerine, J. (1892). Contribution a l'étude anatomopathologique et clinique des différentes variétés de cécité verbale. *Société de Biologie*.
- Di Filippo, G., de Luca, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2006a). Lexicality and Stimulus Length Effects in Italian Dyslexics : Role of the Overadditivity Effect. *Child Neuropsychology*, 12(2), 141-149. <https://doi.org/10.1080/09297040500346571>
- Di Filippo, G., de Luca, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2006b). Lexicality and Stimulus Length Effects in Italian Dyslexics : Role of the Overadditivity Effect. *Child Neuropsychology*, 12(2), 141-149. <https://doi.org/10.1080/09297040500346571>
- Di Folco, C., Guez, A., Peyre, H., & Ramus, F. (2022). Epidemiology of reading disability : A comparison of DSM-5 and ICD-11 criteria. *Scientific Studies of Reading*, 26(4), 337-355. <https://doi.org/10.1080/10888438.2021.1998067>
- Direction de l'Evaluation, de la Prospective et de la Performance. (2021). L'environnement social, économique et familial. In *Géographie de l'École 2021 : N° 13* (Ministère de l'Education Nationale, de la Jeunesse et des Sports).
- Dotan, D., & Friedmann, N. (2015). Steps towards understanding the phonological output buffer and its role in the production of numbers, morphemes, and function words. *Cortex*, 63, 317-351. <https://doi.org/10.1016/j.cortex.2014.08.014>
- Dotan, D., & Friedmann, N. (2019). Separate mechanisms for number reading and word reading : Evidence from selective impairments. *Cortex*, 114, 176-192. <https://doi.org/10.1016/j.cortex.2018.05.010>

- Duranovic, M., Senka, S., & Babic-Gavric, B. (2018). Influence of increased letter spacing and font type on the reading ability of dyslexic children. *Annals of Dyslexia*, 68(3), 218-228. <https://doi.org/10.1007/s11881-018-0164-z>
- Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers : Phonological representations in dyslexia. *Scandinavian Journal of Psychology*, 46(4), 375-384. <https://doi.org/10.1111/j.1467-9450.2005.00468.x>
- Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount : Implications for group differences in response latency. *Psychological Bulletin*, 125(6), 777-799. <https://doi.org/10.1037/0033-2909.125.6.777>
- Faust, M. E., Balota, D. A., Spieler, D. H., Ferraro, F. R., Faust, M. E., Psychology, D. O., Balota, B. D. A., Spieler, D. H., & Department, P. (1999). Individual differences in information-processing rate and amount : Implications for group differences in response latency. *Psychological Bulletin*, 777-799.
- Feng, X., Altarelli, I., Monzalvo, K., Ding, G., Ramus, F., Shu, H., Dehaene, S., Meng, X., & Dehaene-Lambertz, G. (2020a). A universal reading network and its modulation by writing system and reading ability in French and Chinese children. *eLife*, 9, e54591. <https://doi.org/10.7554/eLife.54591>
- Feng, X., Altarelli, I., Monzalvo, K., Ding, G., Ramus, F., Shu, H., Dehaene, S., Meng, X., & Dehaene-Lambertz, G. (2020b). A universal reading network and its modulation by writing system and reading ability in French and Chinese children. *eLife*, 9, e54591. <https://doi.org/10.7554/eLife.54591>
- Feng, X., Monzalvo, K., Dehaene, S., & Dehaene-Lambertz, G. (2022). Evolution of reading and face circuits during the first three years of reading acquisition. *NeuroImage*, 259, 119394. <https://doi.org/10.1016/j.neuroimage.2022.119394>
- Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society*, 31(2), 119-122. <https://doi.org/10.3758/BF03334157>
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French Lexicon Project : Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, 42(2), 488-496. <https://doi.org/10.3758/BRM.42.2.488>
- Fiebach, C. J., Friederici, A. D., Müller, K., & Cramon, D. Y. von. (2002). fMRI Evidence for Dual Routes to the Mental Lexicon in Visual Word Recognition. *Journal of Cognitive Neuroscience*, 14(1), 11-23. <https://doi.org/10.1162/089892902317205285>
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12(6), 627-635. [https://doi.org/10.1016/S0022-5371\(73\)80042-8](https://doi.org/10.1016/S0022-5371(73)80042-8)
- Friedmann, N., Biran, M., & Gvion, A. (2012). Patterns of visual dyslexia : Patterns of visual dyslexia. *Journal of Neuropsychology*, 6(1), 1-30. <https://doi.org/10.1111/j.1748-6653.2011.02000.x>
- Friedmann, N., & Coltheart, M. (2018). 35 .Types of developmental dyslexia. In E. Dattner & D. Ravid (Éds.), *Handbook of Communication Disorders* (p. 721-752). De Gruyter. <https://doi.org/10.1515/9781614514909-036>
- Friedmann, N., Dotan, D., & Rahamim, E. (2010a). Is the visual analyzer orthographic-specific? Reading words and numbers in letter position dyslexia. *Cortex*, 46(8), 982-1004. <https://doi.org/10.1016/j.cortex.2009.08.007>

- Friedmann, N., Dotan, D., & Rahamim, E. (2010b). Is the visual analyzer orthographic-specific? Reading words and numbers in letter position dyslexia. *Cortex*, 46(8), 982-1004. <https://doi.org/10.1016/j.cortex.2009.08.007>
- Friedmann, N., Gvion, A., & Nisim, R. (2015). Insights from letter position dyslexia on morphological decomposition in reading. *Frontiers in Human Neuroscience*, 9, 143. <https://doi.org/10.3389/fnhum.2015.00143>
- Friedmann, N., & Haddad-Hanna, M. (2012a). Letter Position Dyslexia in Arabic : From Form to Position. *Behavioural Neurology*, 25(3), 193-203. <https://doi.org/10.1155/2012/296974>
- Friedmann, N., & Haddad-Hanna, M. (2014a). Types of Developmental Dyslexia in Arabic. In E. Saiegh-Haddad & R. M. Joshi (Éds.), *Handbook of Arabic Literacy* (Vol. 9, p. 119-151). Springer Netherlands. [https://doi.org/10.1007/978-94-017-8545-7\\_6](https://doi.org/10.1007/978-94-017-8545-7_6)
- Friedmann, N., & Haddad-Hanna, M. (2014b). Types of Developmental Dyslexia in Arabic. In E. Saiegh-Haddad & R. M. Joshi (Éds.), *Handbook of Arabic Literacy* (Vol. 9, p. 119-151). Springer Netherlands. [https://doi.org/10.1007/978-94-017-8545-7\\_6](https://doi.org/10.1007/978-94-017-8545-7_6)
- Friedmann, N., & Haddad-Hanna, M. (2014c). Types of Developmental Dyslexia in Arabic. In E. Saiegh-Haddad & R. M. Joshi (Éds.), *Handbook of Arabic Literacy* (Vol. 9, p. 119-151). Springer Netherlands. [https://doi.org/10.1007/978-94-017-8545-7\\_6](https://doi.org/10.1007/978-94-017-8545-7_6)
- Friedmann, N., & Haddad-Hanna, M. (2012b). *Letter Position Dyslexia in Arabic : From Form to Position*. Behavioural Neurology; Hindawi. <https://doi.org/10.3233/BEN-2012-119004>
- Friedmann, N., Kerbel, N., & Shvimer, L. (2010a). Developmental attentional dyslexia. *Cortex*, 46(10), 1216-1237. <https://doi.org/10.1016/j.cortex.2010.06.012>
- Friedmann, N., Kerbel, N., & Shvimer, L. (2010b). Developmental attentional dyslexia. *Cortex*, 46(10), 1216-1237. <https://doi.org/10.1016/j.cortex.2010.06.012>
- Friedmann, N., & Lukov, L. (2008). Developmental surface dyslexias. *Cortex*, 44(9), 1146-1160. <https://doi.org/10.1016/j.cortex.2007.09.005>
- Friedmann, N., & Rahamim, E. (2007a). Developmental letter position dyslexia. *Journal of Neuropsychology*, 1(2), 201-236. <https://doi.org/10.1348/174866407X204227>
- Friedmann, N., & Rahamim, E. (2007b). Developmental letter position dyslexia. *Journal of Neuropsychology*, 1(2), 201-236. <https://doi.org/10.1348/174866407X204227>
- Friedmann, N., & Rahamim, E. (2014a). *What Can Reduce Letter Migrations in Letter Position Dyslexia*. <https://doi.org/10.1111/J.1467-9817.2011.01525.X>
- Friedmann, N., & Rahamim, E. (2014b). What can reduce letter migrations in letter position dyslexia? : TREATMENT OF LETTER POSITION DYSLEXIA. *Journal of Research in Reading*, 37(3), 297-315. <https://doi.org/10.1111/j.1467-9817.2011.01525.x>
- Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001a). Oral Reading Fluency as an Indicator of Reading Competence : A Theoretical, Empirical, and Historical Analysis. In *The Role of Fluency in Reading Competence, Assessment, and Instruction*. Routledge.
- Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001b). Oral Reading Fluency as an Indicator of Reading Competence : A Theoretical, Empirical, and Historical Analysis. *Scientific Studies of Reading*, 5(3), 239-256. [https://doi.org/10.1207/S1532799XSSR0503\\_3](https://doi.org/10.1207/S1532799XSSR0503_3)
- Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., & Cohen, L. (2006). Direct Intracranial, fMRI, and Lesion Evidence for the

- Causal Role of Left Inferotemporal Cortex in Reading. *Neuron*, 50(2), 191-204.  
<https://doi.org/10.1016/j.neuron.2006.03.031>
- Gala, N., Tack, A., Javourey-Drevet, L., François, T., & Ziegler, J. (2020). *Alector : A Parallel Corpus of Simplified French Texts with Alignments of Misreadings by Poor and Dyslexic Readers. Proceedings of the Twelfth Language Resources and Evaluation Conference*, 1353-1361.  
<https://aclanthology.org/2020.lrec-1.169>
- Galliussi, J., Perondi, L., Chia, G., Gerbino, W., & Bernardis, P. (2020). Inter-letter spacing, inter-word spacing, and font with dyslexia-friendly features : Testing text readability in people with and without dyslexia. *Annals of Dyslexia*, 70(1), 141-152. <https://doi.org/10.1007/s11881-020-00194-x>
- Germano, G. D., Reilhac, C., Capellini, S. A., & Valdois, S. (2014). The phonological and visual basis of developmental dyslexia in Brazilian Portuguese reading children. *Frontiers in Psychology*, 5.  
<https://doi.org/10.3389/fpsyg.2014.01169>
- Gori, S., & Facoetti, A. (2014). Perceptual learning as a possible new approach for remediation and prevention of developmental dyslexia. *Vision Research*, 99, 78-87.  
<https://doi.org/10.1016/j.visres.2013.11.011>
- Goswami, U. (2015a). Sensory theories of developmental dyslexia : Three challenges for research. *Nature Reviews Neuroscience*, 16(1), 43-54. <https://doi.org/10.1038/nrn3836>
- Goswami, U. (2015b). Visual attention span deficits and assessing causality in developmental dyslexia. *Nature Reviews Neuroscience*, 16(4), 225-226. <https://doi.org/10.1038/nrn3836-c2>
- Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition*, 123(2), 280-292. <https://doi.org/10.1016/j.cognition.2012.01.003>
- Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition : A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics*, 47(2), 191-198. <https://doi.org/10.3758/BF03205983>
- Guggenheim, R., & Friedmann, N. (2014, juillet). *Previously undescribed type of developmental phonological dyslexia*. In International Neuropsychological Society Meeting.
- Gur, M., & Snodderly, D. M. (1997). A dissociation between brain activity and perception : Chromatically opponent cortical neurons signal chromatic flicker that is not perceived. *Vision Res*, 37(4), 377-382.
- Güven, S., & Friedmann, N. (2019a). Developmental Letter Position Dyslexia in Turkish, a Morphologically Rich and Orthographically Transparent Language. *Frontiers in Psychology*, 10.  
<https://doi.org/10.3389/fpsyg.2019.02401>
- Güven, S., & Friedmann, N. (2019b). Developmental Letter Position Dyslexia in Turkish, a Morphologically Rich and Orthographically Transparent Language. *Frontiers in Psychology*, 10.  
<https://doi.org/10.3389/fpsyg.2019.02401>
- Güven, S., & Friedmann, N. (2021). Vowel dyslexia in Turkish : A window to the complex structure of the sublexical route. *PLOS ONE*, 16(3), e0249016. <https://doi.org/10.1371/journal.pone.0249016>
- Hakvoort, B., Van Den Boer, M., Leenaars, T., Bos, P., & Tijms, J. (2017). Improvements in reading accuracy as a result of increased interletter spacing are not specific to children with dyslexia. *Journal of Experimental Child Psychology*, 164, 101-116. <https://doi.org/10.1016/j.jecp.2017.07.010>
- Heikkilä, R., Aro, M., Närhi, V., Westerholm, J., & Ahonen, T. (2013). Does Training in Syllable Recognition Improve Reading Speed? A Computer-Based Trial With Poor Readers From Second and

- Third Grade. *Scientific Studies of Reading*, 17(6), 398-414.  
<https://doi.org/10.1080/10888438.2012.753452>
- Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005a). Reading Fluency Assessment and Instruction : What, Why, and How? *The Reading Teacher*, 58(8), 702-714. <https://doi.org/10.1598/RT.58.8.1>
- Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005b). Reading Fluency Assessment and Instruction : What, Why, and How? *The Reading Teacher*, 58(8), 702-714. <https://doi.org/10.1598/RT.58.8.1>
- Humphreys, G. W., & Mayall, K. (2001). A peripheral reading deficit under conditions of diffuse visual attention. *Cognitive Neuropsychology*, 18(6), 551-576. <https://doi.org/10.1080/02643290042000242>
- Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S., & Zorman, M. (2010). *BALE (Batterie Analytique du Langage Ecrit)*.
- Jaeger, T. F. (2008). Categorical data analysis : Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434-446.  
<https://doi.org/10.1016/j.jml.2007.11.007>
- Joseph, H., & Powell, D. (2022). Does a specialist typeface affect how fluently children with and without dyslexia process letters, words, and passages? *Dyslexia*, 28(4), 448-470.  
<https://doi.org/10.1002/dys.1727>
- Juphard, A., Carbonnel, S., & Valdois, S. (2004a). Length effect in reading and lexical decision : Evidence from skilled readers and a developmental dyslexic participant. *Brain and cognition*, 55(2), 332-340.
- Juphard, A., Carbonnel, S., & Valdois, S. (2004b). Length effect in reading and lexical decision : Evidence from skilled readers and a developmental dyslexic participant. *Brain and Cognition*, 55(2), 332-340. <https://doi.org/10.1016/j.bandc.2004.02.035>
- Katzir, T., Hershko, S., & Halamish, V. (2013). The Effect of Font Size on Reading Comprehension on Second and Fifth Grade Children : Bigger Is Not Always Better. *PLoS ONE*, 8(9), e74061.  
<https://doi.org/10.1371/journal.pone.0074061>
- Kezilas, Y., Kohnen, S., McKague, M., & Castles, A. (2014). The locus of impairment in English developmental letter position dyslexia. *Frontiers in Human Neuroscience*, 8, 356.  
<https://doi.org/10.3389/fnhum.2014.00356>
- Kohnen, S., Nickels, L., Castles, A., Friedmann, N., & McArthur, G. (2012a). When 'slime' becomes 'smile' : Developmental letter position dyslexia in English. *Neuropsychologia*, 50(14), 3681-3692.  
<https://doi.org/10.1016/j.neuropsychologia.2012.07.016>
- Kohnen, S., Nickels, L., Castles, A., Friedmann, N., & McArthur, G. (2012b). When 'slime' becomes 'smile' : Developmental letter position dyslexia in English. *Neuropsychologia*, 50(14), 3681-3692.  
<https://doi.org/10.1016/j.neuropsychologia.2012.07.016>
- Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-Cabral, L., & Morais, J. (2011). Enantiomorphy through the Looking-Glass : Literacy effects on mirror-image discrimination. *Journal of Experimental Psychology : General*, 140(2), 210-238.
- Koornneef, A., & Kraal, A. (2022). Does BeeLine Reader's gradient-coloured font improve the readability of digital texts for beginning readers? *Computers in Human Behavior Reports*, 6, 100197.  
<https://doi.org/10.1016/j.chbr.2022.100197>
- Krolak-Salmon, P., Henaff, M. A., Tallon-Baudry, C., Yvert, B., Guenot, M., Vighetto, A., Mauguier, F., & Bertrand, O. (2003). Human lateral geniculate nucleus and visual cortex respond to screen flicker. *Ann Neurol*, 53(1), 73-80.



- Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., O'Donovan, M., Williams, J., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., Tóth, D., Honbolygó, F., Csépe, V., Bogliotti, C., Iannuzzi, S., Chaix, Y., Démonet, J.-F., ... Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity : Cross-linguistic predictors of dyslexia. *Journal of Child Psychology and Psychiatry*, 54(6), 686-694. <https://doi.org/10.1111/jcpp.12029>
- Lapeyre, E., Melmi, J.-B., Colé, P., & Calabrèse, A. (2023). *Can pulsed lighting help adults with dyslexia to read better : Testing LeFloch and Ropars's hypothesis* [Preprint]. In Review. <https://doi.org/10.21203/rs.3.rs-3185291/v1>
- Le Floch, A., & Ropars, G. (2017). Left-right asymmetry of the Maxwell spot centroids in adults without and with dyslexia. *Proceedings of the Royal Society B: Biological Sciences*, 284(1865), 20171380. <https://doi.org/10.1098/rspb.2017.1380>
- Lee, J., & Chong, S. C. (2021). Quality of average representation can be enhanced by refined individual items. *Attention, Perception, & Psychophysics*, 83(3), 970-981. <https://doi.org/10.3758/s13414-020-02139-3>
- Lee, K., & Chen, X. (2019). An emergent interaction between reading fluency and vocabulary in the prediction of reading comprehension among French immersion elementary students. *Reading and Writing*, 32(7), 1657-1679. <https://doi.org/10.1007/s11145-018-9920-z>
- Lefavrais, P. (2005). *ALOUETTE-R - Test d'analyse de la lecture et de la dyslexie*.
- Lervag, A., & Hulme, C. (2009). Rapid Automatized Naming (RAN) Taps a Mechanism That Places Constraints on the Development of Early Reading Fluency. *Psychological Science*, 20(8), 1040-1048.
- Lindgren, S. D., De Renzi, E., & Richman, L. C. (1985). Cross-National Comparisons of Developmental Dyslexia in Italy and the United States. *Child Development*, 56(6), 1404. <https://doi.org/10.2307/1130460>
- Livingstone, M. S., Rosen, G. D., Drislane, F. W., & Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Sciences*, 88(18), 7943-7947. <https://doi.org/10.1073/pnas.88.18.7943>
- Lobier, M., Zoubrinetzky, R., & Valdois, S. (2012). The visual attention span deficit in dyslexia is visual and not verbal. *Cortex*, 48(6), 768-773. <https://doi.org/10.1016/j.cortex.2011.09.003>
- Lubineau, M., Watkins, C. P., Glasel, H., & Dehaene, S. (2023). *Why do young readers vary in reading fluency? The impact of word length and frequency in French 6th graders* [Preprint]. Neuroscience. <https://doi.org/10.1101/2023.01.30.526188>
- Lukov, L., Friedmann, N., Shalev, L., Khentov-Kraus, L., Shalev, N., Lorber, R., & Guggenheim, R. (2015a). Dissociations between developmental dyslexias and attention deficits. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01501>
- Lukov, L., Friedmann, N., Shalev, L., Khentov-Kraus, L., Shalev, N., Lorber, R., & Guggenheim, R. (2015b). Dissociations between developmental dyslexias and attention deficits. *Frontiers in Psychology*, 5. <https://www.frontiersin.org/article/10.3389/fpsyg.2014.01501>
- Łuniewska, M., Wójcik, M., & Jednoróg, K. (2022). The effect of inter-letter spacing on reading performance and eye movements in typically reading and dyslexic children. *Learning and Instruction*, 80, 101576. <https://doi.org/10.1016/j.learninstruc.2021.101576>
- Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, A. (1996). On the bases of two subtypes of development dyslexia. *Cognition*, 58(2), 157-195. [https://doi.org/10.1016/0010-0277\(95\)00679-6](https://doi.org/10.1016/0010-0277(95)00679-6)

- Marinus, E., & De Jong, P. F. (2010). Size does not matter, frequency does : Sensitivity to orthographic neighbors in normal and dyslexic readers. *Journal of Experimental Child Psychology, 106*(2-3), 129-144. <https://doi.org/10.1016/j.jecp.2010.01.004>
- Marinus, E., Mostard, M., Segers, E., Schubert, T. M., Madelaine, A., & Wheldall, K. (2016). A Special Font for People with Dyslexia : Does it Work and, if so, why? *Dyslexia, 22*(3), 233-244. <https://doi.org/10.1002/dys.1527>
- Martens, V. E. G., & De Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language, 98*(2), 140-149. <https://doi.org/10.1016/j.bandl.2006.04.003>
- Martens, V. E. G., & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language, 98*(2), 140-149. <https://doi.org/10.1016/j.bandl.2006.04.003>
- Masri, R. A., Grünert, U., & Martin, P. R. (2020). Analysis of Parvocellular and Magnocellular Visual Pathways in Human Retina. *The Journal of Neuroscience, 40*(42), 8132-8148. <https://doi.org/10.1523/JNEUROSCI.1671-20.2020>
- McCann, R. S., & Besner, D. (1987). Reading Pseudohomophones : Implications for Models of Pronunciation Assembly and the Locus of Word-Frequency Effects in Naming. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 14-24.
- McCloskey, M., & Rapp, B. (2000a). A Visually Based Developmental Reading Deficit. *Journal of Memory and Language, 43*(2), 157-181. <https://doi.org/10.1006/jmla.2000.2724>
- McCloskey, M., Rapp, B., Yantis, S., Rubin, G., Bacon, W. F., Dagnelie, G., Gordon, B., Aliminosa, D., Boatman, D. F., Badecker, W., Johnson, D. N., Tusa, R. J., & Palmer, E. (1995). A developmental deficit in localizing objects from vision. *Psychological Science, 6*(2).
- McCloskey, & Rapp. (2000b). A visually based developmental reading deficit. *Journal of memory and language, 43*(2), 157-181.
- Monzalvo, K. (2011). *Etude chez l'enfant normal et dyslexique de l'impact sur les réseaux corticaux visuel et linguistique d'une activité culturelle : La lecture*. Paris 6.
- Monzalvo, K., & Dehaene-Lambertz, G. (2013). How reading acquisition changes children's spoken language network. *Brain and Language, 127*(3), 356-365. <https://doi.org/10.1016/j.bandl.2013.10.009>
- Nakamura, K., Kuo, W.-J., Pegado, F., Cohen, L., Tzeng, O. J. L., & Dehaene, S. (2012). Universal brain systems for recognizing word shapes and handwriting gestures during reading. *Proceedings of the National Academy of Sciences, 109*(50), 20762-20767. <https://doi.org/10.1073/pnas.1217749109>
- National Reading Panel. (2000). *Teaching children to read : An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)*. National Institute of Child Health & Development. <https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf>
- Nese, J. F. T., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., & Tindal, G. (2013). In search of average growth : Describing within-year oral reading fluency growth across Grades 1–8. *Journal of School Psychology, 51*(5), 625-642. <https://doi.org/10.1016/j.jsp.2013.05.006>
- New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006a). Reexamining the word length effect in visual word recognition : New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13*(1), 45-52. <https://doi.org/10.3758/BF03193811>

- New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006b). Reexamining the word length effect in visual word recognition : New evidence from the English Lexicon Project. *Psychon Bull Rev*, 13(1), 45-52.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2 : A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516-524. <https://doi.org/10.3758/BF03195598>
- Norris, D. (2006). The Bayesian reader : Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, 113(2), 327-357. <https://doi.org/10.1037/0033-295X.113.2.327>
- O'Brien, B. A., Mansfield, J. S., & Legge, G. E. (2005). The effect of print size on reading speed in dyslexia. *Journal of Research in Reading*, 28(3), 332-349. <https://doi.org/10.1111/j.1467-9817.2005.00273.x>
- Orton, S. (1925). "Word-blindness" in school children. *Archives of Neurology and Psychiatry*, 14, 581-615.
- Orton, S. (1928). A physiological theory of reading disability and stuttering in children. *The New England Journal of Medicine*, 199, 1046-1052.
- Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound : Still a good horse race. *Psychological Research*, 53(1), 13-24. <https://doi.org/10.1007/BF00867328>
- Paterson, K. B., Read, J., McGowan, V. A., & Jordan, T. R. (2015). Children and adults both see 'pirates' in 'parties' : Letter-position effects for developing readers and skilled adult readers. *Developmental Science*, 18(2), 335-343. <https://doi.org/10.1111/desc.12222>
- Pegado, F., Nakamura, K., Braga, L. W., Ventura, P., Nunes Filho, G., Pallier, C., Jobert, A., Morais, J., Cohen, L., Kolinsky, R., & Dehaene, S. (2014). Literacy breaks mirror invariance for visual stimuli : A behavioral study with adult illiterates. *Journal of Experimental Psychology. General*, 143(2), 887-894. <https://doi.org/10.1037/a0033198>
- Pegado, F., Nakamura, K., Cohen, L., & Dehaene, S. (2011a). Breaking the symmetry : Mirror discrimination for single letters but not for pictures in the Visual Word Form Area. *NeuroImage*, 55(2), 742-749. <https://doi.org/10.1016/j.neuroimage.2010.11.043>
- Pegado, F., Nakamura, K., Cohen, L., & Dehaene, S. (2011b). Breaking the symmetry : Mirror discrimination for single letters but not for pictures in the Visual Word Form Area. *NeuroImage*, 55(2), 742-749. <https://doi.org/10.1016/j.neuroimage.2010.11.043>
- Pflugshaupt, T., Nyffeler, T., Vonwartburg, R., Wurtz, P., Luthi, M., Hubl, D., Gutbrod, K., Juengling, F., Hess, C., & Muri, R. (2007). When left becomes right and vice versa : Mirrored vision after cerebral hypoxia. *Neuropsychologia*, 45(9), 2078-2091. <https://doi.org/10.1016/j.neuropsychologia.2007.01.018>
- Pinna, B., & Deiana, K. (2018). On the Role of Color in Reading and Comprehension Tasks in Dyslexic Children and Adults. *I-Perception*, 9(3), 204166951877909. <https://doi.org/10.1177/2041669518779098>
- Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). *Listening to children read aloud : Oral fluency*. Department of Education, National Center for Education Statistics.
- Potier Watkins, C., Dehaene, S., & Friedmann, N. (2023). *Characterizing Different Types of Developmental Dyslexias in French : The Malabi Screener*. <https://doi.org/10.23668/PSYCHARCHIVES.12512>

- Potier-Watkins, C., Dehaene, S., & Dehaene, O. (2019). Automatic construction of a phonics curriculum for reading education using the Transformer neural network. *International Conference on Artificial Intelligence in Education*, 226-231.
- Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The Nonword Reading Deficit in Developmental Dyslexia : A Review. *Reading Research Quarterly*, 27(1), 28. <https://doi.org/10.2307/747832>
- Ramus, F. (2001). Talk of two theories. *Nature*, 412(6845), 393-394. <https://doi.org/10.1038/35086683>
- Ramus, F. (2003a). Developmental dyslexia : Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, 13(2), 212-218. [https://doi.org/10.1016/S0959-4388\(03\)00035-7](https://doi.org/10.1016/S0959-4388(03)00035-7)
- Ramus, F. (2003b). Theories of developmental dyslexia : Insights from a multiple case study of dyslexic adults. *Brain*, 126(4), 841-865. <https://doi.org/10.1093/brain/awg076>
- Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., & Scotto Di Covella, L. (2018). Neuroanatomy of developmental dyslexia : Pitfalls and promise. *Neuroscience & Biobehavioral Reviews*, 84, 434-452. <https://doi.org/10.1016/j.neubiorev.2017.08.001>
- Ramus, F., Pidgeon, E., & Frith, U. (2003). The relationship between motor control and phonology in dyslexic children : Motor control and phonology in dyslexic children. *Journal of Child Psychology and Psychiatry*, 44(5), 712-722. <https://doi.org/10.1111/1469-7610.00157>
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). A Diffusion Model Account of the Lexical Decision Task. *Psychological review*, 111(1), 159-182. <https://doi.org/10.1037/0033-295X.111.1.159>
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading : Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14(3), 191-201. <https://doi.org/10.3758/BF03197692>
- Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, 16(8), 829-837. [https://doi.org/10.1016/0042-6989\(76\)90143-7](https://doi.org/10.1016/0042-6989(76)90143-7)
- Rayner, K., Murphy, L. A., Henderson, J. M., & Pollatsek, A. (1989a). Selective attentional dyslexia. *Cognitive Neuropsychology*, 6(4), 357-378. <https://doi.org/10.1080/02643298908253288>
- Rayner, K., Murphy, L. A., Henderson, J. M., & Pollatsek, A. (1989b). Selective attentional dyslexia. *Cognitive Neuropsychology*, 6(4), 357-378.
- Rello, L., & Baeza-Yates, R. (2013). Good fonts for dyslexia. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, 1-8. <https://doi.org/10.1145/2513383.2513447>
- Rello, L., & Baeza-Yates, R. (2017). How to present more readable text for people with dyslexia. *Universal Access in the Information Society*, 16(1), 29-49. <https://doi.org/10.1007/s10209-015-0438-8>
- Roongpraiwan, R., Ruangdaraganon, N., Visudhiphan, P., & Santikul, K. (2002). Prevalence and clinical characteristics of dyslexia in primary school students. *Journal of the Medical Association of Thailand = Chotmaihet Thangphaet*, 85 Suppl 4, S1097-1103.
- Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W.-J., Bick, A., Frost, S. J., Hancock, R., Wu, D. H., Mencl, W. E., Duñabeitia, J. A., Lee, J.-R., Oliver, M., Zevin, J. D., Hoeft, F., Carreiras, M., Tzeng, O. J. L., Pugh, K. R., & Frost, R. (2015). Universal brain signature of proficient reading : Evidence from four contrasting languages. *Proceedings of the National Academy of Sciences*, 112(50), 15510-15515. <https://doi.org/10.1073/pnas.1509321112>

- Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-Assisted Remedial Reading Intervention for School Beginners at Risk for Reading Disability : Computer-Assisted Reading Intervention. *Child Development, 82*(3), 1013-1028. <https://doi.org/10.1111/j.1467-8624.2011.01580.x>
- Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J.-F., Bricout, L., Billard, C., Nguyen-Morel, M.-A., Le Heuzey, M.-F., Soares-Boucaud, I., George, F., Ziegler, J. C., & Ramus, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology, 52*(10), 1503-1516. <https://doi.org/10.1037/dev0000184>
- Schmalz, X., Marinus, E., & Castles, A. (2013a). Phonological decoding or direct access? Regularity effects in lexical decisions of Grade 3 and 4 children. *Quarterly Journal of Experimental Psychology, 66*(2), 338-346. <https://doi.org/10.1080/17470218.2012.711843>
- Schmalz, X., Marinus, E., & Castles, A. (2013b). Phonological decoding or direct access? Regularity effects in lexical decisions of Grade 3 and 4 children. *Quarterly Journal of Experimental Psychology, 66*(2), 338-346. <https://doi.org/10.1080/17470218.2012.711843>
- Schröter, P., & Schroeder, S. (2017). The Developmental Lexicon Project : A behavioral database to investigate visual word recognition across the lifespan. *Behavior Research Methods, 49*(6), 2183-2203. <https://doi.org/10.3758/s13428-016-0851-9>
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568. <https://doi.org/10.1037/0033-295X.96.4.523>
- Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23*(3), 383-404. [https://doi.org/10.1016/S0022-5371\(84\)90270-6](https://doi.org/10.1016/S0022-5371(84)90270-6)
- Sela, I., Izzetoglu, M., Izzetoglu, K., & Onaral, B. (2014). A Functional Near-Infrared Spectroscopy Study of Lexical Decision Task Supports the Dual Route Model and the Phonological Deficit Theory of Dyslexia. *Journal of Learning Disabilities, 47*(3), 279-288. <https://doi.org/10.1177/0022219412451998>
- Seymour, P. H. K., Aro, M., Erskine, J. M., & collaboration with COST Action A8 network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143-174. <https://doi.org/10.1348/000712603321661859>
- Shallice, T., & Warrington, E. K. (1977). The possible role of selective attention in acquired dyslexia. *Neuropsychologia, 15*(1), 31-41. [https://doi.org/10.1016/0028-3932\(77\)90112-9](https://doi.org/10.1016/0028-3932(77)90112-9)
- Share, D. L. (1995). Phonological recoding and self-teaching : Sine qua non of reading acquisition. *Cognition, 55*(2), 151-218. [https://doi.org/10.1016/0010-0277\(94\)00645-2](https://doi.org/10.1016/0010-0277(94)00645-2)
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence That Dyslexia May Represent the Lower Tail of a Normal Distribution of Reading Ability. *New England Journal of Medicine, 326*(3), 145-150. <https://doi.org/10.1056/NEJM199201163260301>
- Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task : A characterization of the mind's bottleneck. *PLoS Biol, 3*(2), e37.
- Sjoblom, A. M., Eaton, E., & Stagg, S. D. (2016). The effects of letter spacing and coloured overlays on reading speed and accuracy in adult dyslexia. *British Journal of Educational Psychology, 86*(4), 630-639. <https://doi.org/10.1111/bjep.12127>

- Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length Effect in Word Naming in Reading : Role of Reading Experience and Reading Deficit in Italian Readers. *Developmental Neuropsychology*, 27(2), 217-235. [https://doi.org/10.1207/s15326942dn2702\\_2](https://doi.org/10.1207/s15326942dn2702_2)
- Sprenger-Charolles, L., Colé, P., Lacert, P., & Serniclaes, W. (2000). On subtypes of developmental dyslexia : Evidence from processing time and accuracy scores. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 54(2), 87-104. <https://doi.org/10.1037/h0087332>
- Sprenger-Charolles, L., Siegel, L. S., Jiménez, J. E., & Ziegler, J. C. (2011). Prevalence and Reliability of Phonological, Surface, and Mixed Profiles in Dyslexia : A Review of Studies Conducted in Languages Varying in Orthographic Depth. *Scientific Studies of Reading*, 15(6), 498-521. <https://doi.org/10.1080/10888438.2010.524463>
- Stagg, S. D., & Kiss, N. (2021). Room to read : The effect of extra-large letter spacing and coloured overlays on reading speed and accuracy in adolescents with dyslexia. *Research in Developmental Disabilities*, 119, 104065. <https://doi.org/10.1016/j.ridd.2021.104065>
- Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology*, 89(1), 114-127. <https://doi.org/10.1037/0022-0663.89.1.114>
- Stefanou, A. (2022). L'éducation Prioritaire. *DEPP, Ministère de l'Éducation Nationale, de la Jeunesse et des Sports*.
- Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neurosciences*, 20(4), 147-152. [https://doi.org/10.1016/S0166-2236\(96\)01005-3](https://doi.org/10.1016/S0166-2236(96)01005-3)
- Temple, C. M. (2003). Deep dyslexia in Williams syndrome. *Journal of Neurolinguistics*, 16(6), 457-488. [https://doi.org/10.1016/S0911-6044\(01\)00044-6](https://doi.org/10.1016/S0911-6044(01)00044-6)
- Temple, C. M. (2006). Developmental and Acquired Dyslexias. *Cortex*, 42(6), 898-910. [https://doi.org/10.1016/S0010-9452\(08\)70434-9](https://doi.org/10.1016/S0010-9452(08)70434-9)
- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. E. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation : Evidence from functional MRI. *Proceedings of the National Academy of Sciences*, 100(5), 2860-2865. <https://doi.org/10.1073/pnas.0030098100>
- Toledano, L., & Friedmann, N. (2023). Letter Migrations between Words in Reading Aloud Can Result either from an Impairment in Orthographic Input or in Phonological Output. *Brain Sciences*, 13(4), 588. <https://doi.org/10.3390/brainsci13040588>
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive Remedial Instruction for Children with Severe Reading Disabilities : Immediate and Long-term Outcomes From Two Instructional Approaches. *Journal of Learning Disabilities*, 34(1), 33-58. <https://doi.org/10.1177/002221940103400104>
- Traficante, D., Luzzatti, C., & Friedmann, N. (2021). *A Complex View of the Grapheme-to-Phoneme Conversion (GPC) Procedure : Evidence for Vowel Developmental Dyslexia from a Shallow Orthography Language*. EasyChair preprint.
- Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faisca, L., Reis, A., & Blomert, L. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of Educational Psychology*, 102(4), 827-842. <https://doi.org/10.1037/a0019465>

- Valdois, S. (2014). What do we understand by attentional disorder in the context of dyslexia? *ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant*, 26(128), 27-35.
- Valdois, S. (2022). The visual-attention span deficit in developmental dyslexia : Review of evidence for a visual-attention-based deficit. *Dyslexia*, 28(4), 397-415. <https://doi.org/10.1002/dys.1724>
- Valdois, S., Bidet-Ildei, C., Lassus-Sangosse, D., Reilhac, C., N'guyen-Morel, M.-A., Guinet, E., & Orliaguet, J.-P. (2011). A visual processing but no phonological disorder in a child with mixed dyslexia. *Cortex*, 47(10), 1197-1218. <https://doi.org/10.1016/j.cortex.2011.05.011>
- Valdois, S., Bosse, M.-L., Ans, B., Carbonnel, S., Zorman, M., David, D., & Pellat, J. (2003). Phonological and visual processing deficits can dissociate in developmental dyslexia : Evidence from two case studies. *Reading and Writing*, 16(6), 541-572. <https://doi.org/10.1023/A:1025501406971>
- Valdois, S., Bosse, M.-L., & Tainturier, M.-J. (2004). The cognitive deficits responsible for developmental dyslexia : Review of evidence for a selective visual attentional disorder. *Dyslexia*, 10(4), 339-363. <https://doi.org/10.1002/dys.284>
- Valdois, S., Lassus-Sangosse, D., & Lobier, M. (2012). The Visual Nature of the Visual Attention Span Disorder in Developmental Dyslexia. In *Visual aspects of dyslexia* (p. 111-122).
- Valdois, S., Peyrin, C., Lassus-Sangosse, D., Lallier, M., Démonet, J.-F., & Kandel, S. (2014). Dyslexia in a French-Spanish bilingual girl : Behavioural and neural modulations following a visual attention span intervention. *Cortex*, 53, 120-145. <https://doi.org/10.1016/j.cortex.2013.11.006>
- Valdois, S., Reilhac, C., Ginestet, E., & Line Bosse, M. (2021). Varieties of Cognitive Profiles in Poor Readers : Evidence for a VAS-Impaired Subtype. *Journal of Learning Disabilities*, 54(3), 221-233. <https://doi.org/10.1177/0022219420961332>
- Vannuscorps, G., Galaburda, A., & Caramazza, A. (2022). Shape-centered representations of bounded regions of space mediate the perception of objects. *Cognitive Neuropsychology*, 39(1-2), 1-50. <https://doi.org/10.1080/02643294.2021.1960495>
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia) : What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45(1), 2-40. <https://doi.org/10.1046/j.0021-9630.2003.00305.x>
- Vidyasagar, T. R. (1999). A neuronal model of attentional spotlight : Parietal guiding the temporal. *Brain Research Reviews*, 30(1), 66-76. [https://doi.org/10.1016/S0165-0173\(99\)00005-3](https://doi.org/10.1016/S0165-0173(99)00005-3)
- Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia : A deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences*, 14(2), 57-63. <https://doi.org/10.1016/j.tics.2009.12.003>
- Weekes, B. (1997). Differential Effects of Number of Letters on Word and Nonword Naming Latency. *Quarterly Journal of Experimental Psychology Section A-human Experimental Psychology - QUART J EXP PSYCH A-HUM EXP P*, 50, 439-456. <https://doi.org/10.1080/027249897392170>
- Weekes, B. S. (1997). Differential Effects of Number of Letters on Word and Nonword Naming Latency. *The Quarterly Journal of Experimental Psychology A*, 50(2), 439-456. <https://doi.org/10.1080/027249897392170>
- Wery, J. J., & Diliberto, J. A. (2017). The effect of a specialized dyslexia font, OpenDyslexic, on reading rate and accuracy. *Annals of Dyslexia*, 67(2), 114-127. <https://doi.org/10.1007/s11881-016-0127-1>
- White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., & Ramus, F. (2006). The role of sensorimotor impairments in dyslexia : A multiple case study of dyslexic children. *Developmental Science*, 9(3), 237-255. <https://doi.org/10.1111/j.1467-7687.2006.00483.x>

- Williams, M. J., Stuart, G. W., Castles, A., & McAnally, K. I. (2003). Contrast sensitivity in subgroups of developmental dyslexia. *Vision Research*, *43*(4), 467-477. [https://doi.org/10.1016/S0042-6989\(02\)00573-4](https://doi.org/10.1016/S0042-6989(02)00573-4)
- Yamada, J. (1995). Developmental Deep Dyslexia in Japanese : A Case Study. *Brain and Language*, *51*(3), 444-457. <https://doi.org/10.1006/brln.1995.1070>
- Yang, L., Li, C., Li, X., Zhai, M., An, Q., Zhang, Y., Zhao, J., & Weng, X. (2022). Prevalence of Developmental Dyslexia in Primary School Children : A Systematic Review and Meta-Analysis. *Brain Sciences*, *12*(2), 240. <https://doi.org/10.3390/brainsci12020240>
- Yeari, M., Isser, M., & Schiff, R. (2017). Do dyslexic individuals present a reduced visual attention span? Evidence from visual recognition tasks of non-verbal multi-character arrays. *Annals of Dyslexia*, *67*(2), 128-146. <https://doi.org/10.1007/s11881-016-0132-4>
- Yeatman, J. D., Tang, K. A., Donnelly, P. M., Yablonski, M., Ramamurthy, M., Karipidis, I. I., Caffarra, S., Takada, M. E., Kanopka, K., Ben-Shachar, M., & Domingue, B. W. (2021). Rapid online assessment of reading ability. *Scientific Reports*, *11*(1), Article 1. <https://doi.org/10.1038/s41598-021-85907-x>
- Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic Depth and Its Impact on Universal Predictors of Reading : A Cross-Language Investigation. *Psychological Science*, *21*(4), 551-559. <https://doi.org/10.1177/0956797610363406>
- Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F.-X., & Perry, C. (2008). Developmental dyslexia and the dual route model of reading : Simulating individual differences and subtypes. *Cognition*, *107*(1), 151-178. <https://doi.org/10.1016/j.cognition.2007.09.004>
- Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages : A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3-29. <https://doi.org/10.1037/0033-2909.131.1.3>
- Ziegler, J. C., Pech-Georgel, C., Dufau, S., & Grainger, J. (2009). Rapid processing of letters, digits and symbols : What purely visual-attentional deficit in developmental dyslexia?: Processing of letters, digits and symbols in dyslexia. *Developmental Science*, *13*(4), F8-F14. <https://doi.org/10.1111/j.1467-7687.2010.00983.x>
- Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages : Language-specific or universal? *Journal of Experimental Child Psychology*, *86*(3), 169-193. [https://doi.org/10.1016/S0022-0965\(03\)00139-5](https://doi.org/10.1016/S0022-0965(03)00139-5)
- Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching : Implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120397. <https://doi.org/10.1098/rstb.2012.0397>
- Ziegler, J. C., Perry, C., & Zorzi, M. (2020). Learning to Read and Dyslexia : From Theory to Intervention Through Personalized Computational Models. *Current Directions in Psychological Science*, *29*(3), 293-300. <https://doi.org/10.1177/0963721420915873>
- Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., & Martelli, M. (2009a). Reading development in an orthographically regular language : Effects of length, frequency, lexicality and global processing ability. *Reading and Writing*, *22*(9), 1053-1079. <https://doi.org/10.1007/s11145-008-9144-8>
- Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., & Martelli, M. (2009b). Reading development in an orthographically regular language : Effects of length, frequency, lexicality and global processing ability. *Reading and Writing*, *22*(9), 1053-1079. <https://doi.org/10.1007/s11145-008-9144-8>



Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005a). Word length effect in early reading and in developmental dyslexia. *Brain and Language*, *93*(3), 369-373. <https://doi.org/10.1016/j.bandl.2004.10.010>

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005b). Word length effect in early reading and in developmental dyslexia. *Brain and Language*, *93*(3), 369-373. <https://doi.org/10.1016/j.bandl.2004.10.010>

Zoccolotti, P., De Luca, M., Judica, A., & Spinelli, D. (2008). Isolating global and specific factors in developmental dyslexia : A study based on the rate and amount model (RAM). *Experimental brain research*, *186*(4), 551-560.

Zoccolotti, P., & Friedmann, N. (2010). From dyslexia to dyslexias, from dysgraphia to dysgraphias, from a cause to causes : A look at current research on developmental dyslexia and dysgraphia. *Cortex*, *46*(10), 1211-1215. <https://doi.org/10.1016/j.cortex.2010.09.003>

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., Bravar, L., George, F., Pech-Georgel, C., & Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences*, *109*(28), 11455-11459. <https://doi.org/10.1073/pnas.1205566109>

Zoubinetzky, R., Bielle, F., & Valdois, S. (2014). New Insights on Developmental Dyslexia Subtypes : Heterogeneity of Mixed Reading Profiles. *PLoS ONE*, *9*(6), e99337. <https://doi.org/10.1371/journal.pone.0099337>

Zoubinetzky, R., Collet, G., Nguyen-Morel, M.-A., Valdois, S., & Serniclaes, W. (2019). Remediation of Allophonic Perception and Visual Attention Span in Developmental Dyslexia : A Joint Assay. *Frontiers in Psychology*, *10*, 1502. <https://doi.org/10.3389/fpsyg.2019.01502>

Zoubinetzky, R., Collet, G., Serniclaes, W., Nguyen-Morel, M.-A., & Valdois, S. (2016). Relationships between Categorical Perception of Phonemes, Phoneme Awareness, and Visual Attention Span in Developmental Dyslexia. *PLOS ONE*, *11*(3), e0151015. <https://doi.org/10.1371/journal.pone.0151015>

# APPENDIX

## Appendix 1. La Mariette

Nous sommes sur une planète étrange. Sur la mare ou dans les bois, partout où coule notre regard, le temps est flache. Une plante gradée pose son parfum dans la nuit. Près de la truche, un enfant brise des galets en dix places fiables. Depuis hier, l'arbi le regarde sur son vélo volé. Un second sur le navire crie à sept chats oviles:

- Cape page dira dure !

Les chats ont compris et vont chez la sorcière.

Cette femme habitait dans une ferme avec une petite crique. Un gentil mantou lui servait de lange. Elle pilait une potion de gaises de limaces et de dents de sardines avec une trace d'allumette. Elle clamait sa colère en disant:

- Klapin vente perte lame rime !

Les chats lui délivrèrent la fiole de tabac. Elle ajouta le comprimé et une carie blanche. Le son d'un orchestre retentit comme un aspect. Elle donna du blacon aux chats qui repartirent.

Les voici au pied du dolmen rogné par les vagues. Dans l'eau près de l'enfer, le dauphin acrobate joue en criant:

- balle selle, cousin !

- dont vent sois puis, répond la dorade grusée d'un air ibole.

Les chats arrivèrent au galop devant l'aquarium et les poissons patrons prirent la fuite dans une dragée de couleurs. Les chats firent un rond en arrière vers leur rabio puis décidèrent de regagner la montagne avec respect.

Sur le chemin ils explorèrent les éballons autour de l'étang que la cure du fleuve avait formé. Une chorale chantait:

- Marcherions paire foire loveau mont sent.

En aval du glabe le patient et la larve au grand cœur mangeaient leurs lignes avec quelques oignons. Les chats passèrent devant différents repratis aux écritures nettes. Comme ce n'étaient pas des escrocs ils aidèrent Marie la maire. Ce fut une belle journée.