Cognitive Neuropsychology

ISSN: 0264-3294 (Print) 1464-0627 (Online) Journal homepage: www.tandfonline.com/journals/pcgn20

Mariette: A screening test for reading errors in primary school

Marie Lubineau, Stanislas Dehaene, Hervé Glasel & Cassandra Potier Watkins

To cite this article: Marie Lubineau, Stanislas Dehaene, Hervé Glasel & Cassandra Potier Watkins (02 Sep 2025): Mariette: A screening test for reading errors in primary school, Cognitive Neuropsychology, DOI: 10.1080/02643294.2025.2544620

To link to this article: https://doi.org/10.1080/02643294.2025.2544620

Published online: 02 Sep 2025.

COGNITIVE NEUROPSYCHOLOGY
https://doi.org/10.1080/02643294.2025.2544620

Mariette: A screening test for reading errors in primary school

Marie Lubineau [a,b], Stanislas Dehaene[a,c], Hervé Glasel[b] and Cassandra Potier Watkins [a,c]

[a]Unité de Neuroimagerie cognitive, CEA, INSERM, Université Paris-Saclay, NeuroSpin center, Saclay, France; [b]Centre de référence pour l'évaluation neuropsychologique de l'enfant (CERENE), Paris, France; [c]Collège de France, Paris, France

ABSTRACT

Dyslexia is a multifaceted condition with diverse manifestations, yet assessment tools too often target limited subtypes, creating diagnostic gaps. This study examines the progression of dyslexia-related reading errors across primary school in typically developing readers, using the Mariette, a French nonsense-text reading screener. Analysis of 812 French children (grades 1–5) revealed systematic decreases in error rates with age, following distinct developmental trajectories. Regularizations of irregular words, misapplication of contextual rules and misreading of digraphs predominated in early grades, while voicing errors nearly disappeared by Grade 2. Clinical testing of the Mariette with 18 struggling readers identified specific reading errors overlooked by standard dyslexia assessments. These findings demonstrate the value of precise error analysis for understanding developmental reading patterns and tailoring targeted educational interventions. By comparing typical and clinical populations, this research advances our understanding of dyslexia's cognitive mechanisms while advocating for more comprehensive diagnostic approaches.

ARTICLE HISTORY
Received 14 January 2025
Revised 29 July 2025
Accepted 1 August 2025

KEYWORDS
Developmental dyslexia; dyslexia screener; children; reading errors; letter position dyslexia; attentional dyslexia; surface dyslexia; phonological buffer dyslexia

Introduction

Developmental dyslexia, a common learning disability, affects between 3% and 12% of the population, depending on language and diagnostic criteria (Di Folco et al., 2022; *Diagnostic and Statistical Manual of Mental Disorders | Psychiatry Online*, s. d.; Lindgren et al., 1985). Given reading's pivotal role in academic success, early detection and intervention are important (McGee et al., 2002). However, diagnosing dyslexia as a deviation from typical reading development is challenging due to three key factors. First, diverse cognitive and neural processes are involved in reading acquisition (Monzalvo & Dehaene-Lambertz, 2013). Second, children develop reading skills at different rates, making it difficult to determine when errors fall outside the expected developmental trajectories. Third, growing evidence suggests that dyslexia can manifest through multiple distinct reading deficits, each tied to different cognitive processes, resulting in varying degrees of reading difficulty and patterns of errors depending on context (Castles & Coltheart, 1993; Coltheart & Kohnen, 2012; Ellis, 1979; Friedmann & Coltheart, 2018; Friedmann & Haddad-Hanna, 2014; Temple, 2006). Despite growing recognition of dyslexia's multi-deficit nature, comprehensive screening tools that account for the developmental variability while providing detailed insights into dyslexia subtypes remain rare in clinical practice. Developing such tools is essential for enabling early and accurate diagnoses, leading to more effective interventions.

Understanding the complexity of reading development and the potential for selective deficits can be aided by the Dual Route Model of reading (Coltheart et al., 2001). Research into the various types of errors made by individuals with dyslexia has been crucial in refining this model and explaining how specific deficits in reading processes can lead to distinct forms of dyslexia. (Castles, 2006; Friedmann & Coltheart, 2018; Potier Watkins et al., 2023). Below, we provide a brief overview of the cognitive processes described by the Dual Route Model, and different subtypes of dyslexia that may arise from selective deficits. For a more detailed discussion, see Friedmann and Coltheart (2018) or Potier Watkins et al. (2023).

According to the Dual Route Model by Friedmann and Coltheart depicted in Figure 1, reading begins with visual analysis, which encompasses several critical processes: identifying abstract letter identities; encoding the position of letters within words; and anchoring letters to their respective words (Friedmann & Coltheart, 2018). Visual processing deficits in dyslexia can manifest in multiple ways, affecting different aspects of this initial analysis stage.

Letter identity dyslexia represents a form of visual processing deficit characterized by difficulty in correctly

CONTACT Marie Lubineau marie.lubineau@college-de-france.fr

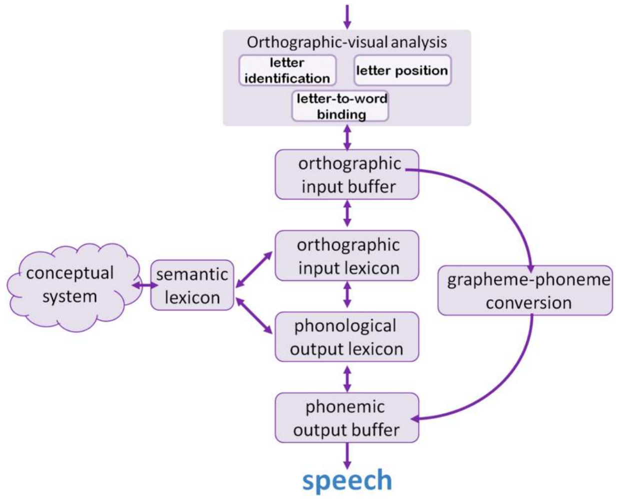

Figure 1. The Dual Route Model for word reading (taken from Friedmann & Coltheart, 2018).

recognizing the abstract of identities of letters from letter shapes (e.g., identifying that B and b correspond to the same letter). Individuals with this rare condition struggle to name, identify, or match upper- and lowercase letters (Brunsdon et al., 2006). Because letter identity dyslexia leads to severe reading impairments involving letter confusion, it is typically diagnosed by the presence of numerous letter substitutions, omissions, and additions. Such cases must be followed up with specific assessments to confirm whether letter identification abilities are intact. Letter-position dyslexia emerges from a disruption in encoding the relative positions of letters, which causes transpositions of letters, such as reading *destiny* as *density* – a phenomenon well-documented across multiple languages (Friedmann & Gvion, 2001; Friedmann & Haddad-Hanna, 2012; Friedmann & Rahamim, 2007; Güven & Friedmann, 2019; Kohnen et al., 2012; Potier Watkins et al., 2023). People with this subtype of dyslexia have also been reported to repeat or omit a letter repeated in a word (Friedmann & Rahamim, 2007; Potier Watkins et al., 2023). The hypothesis is that when letter-position encoding is impaired, distinguishing the position of each letter becomes difficult, leading to a letter showing up in two different positions or omitted in one position. Attentional dyslexia describes a deficit in binding letters to the correct word: letters migrate from one word to another, leading to errors such as reading *win fed* as *fin fed* (Friedmann et al., 2010; Potier Watkins et al., 2023; Rayner et al., 1989). In the vast majority of errors, letters that migrate maintain their relative position in the new word, highlighting that letter position coding within a word is intact in attentional dyslexia, and that this type of dyslexia therefore differs sharply from the previous one (Friedmann et al., 2010). Importantly, attentional migrations tend to occur equally between the first and second word in word-pair reading (Potier Watkins et al., 2023; Shallice & Warrington, 1977), demonstrating that the migration is not due to phonological perseveration, but rather to impaired orthographic processing.

Orthographic-visual analyzer dyslexia (also referred to as *visual dyslexia*) is characterized by reading errors that involve substitutions, omissions, or additions of letters that share visual similarity with the target word. For example, "unicorn" may be misread as "united" or "acorn" (Friedmann et al., 2012). These errors cannot only be explained by difficulties with letter transpositions, migrations between words, or basic letter identification deficits, as single-letter recognition remains intact. The hypothesis is that this dyslexia reflects a disruption of the whole visual-orthographic analysis stage. Because this stage is disrupted, readers produce a wide and variable range of errors (migrations, substitutions, additions, omissions) that superficially resemble those seen in letter-identity, letter-position, or attentional dyslexia, yet do not conform to the specific patterns that define those disorders. In other words, orthographic-visual analyzer dyslexia is distinguished by the presence of diverse, non-systematic visual errors despite intact basic letter recognition (Friedmann et al., 2012; Friedmann & Coltheart, 2018). This is important to distinguish from letter identity dyslexia, as both disorders can cause similar reading errors. However, in the case of letter identity dyslexia, the ability to recognize single letters is impaired. Furthermore, when errors of addition, omission or substitution occur solely on one side of the word - either at the beginning or the end – these errors point to the presence of neglexia (shorthand for neglect dyslexia) (Ellis et al., 1987; Friedmann & Nachman-Katz, 2004; Nachman-Katz & Friedmann, 2010). Orthographic-visual analyzer dyslexia should also be distinguished from a deficit in the orthographic input buffer. The orthographic input buffer functions as a short-term storage system that temporarily maintains orthographic information until it can be fully processed by downstream components. Although deficits in this buffer can produce error types similar to those seen in orthographic-visual analyzer dyslexia, a key distinguishing feature is their sensitivity to word length: errors tend to occur more frequently as word length increases. When such length-dependent error patterns are absent, the disruption is more likely to stem from the orthographic-visual analysis stage rather than from the buffer itself (Friedmann & Coltheart, 2018).

From this visual stage, words are pieced in an orthographic input buffer, from which reading then flows to two separate but parallel routes in the Dual Route

Model. Most reading, in the expert reader, is done through a rapid lexical route that quickly identifies words stored in the mental lexicon and accesses their meaning (Coltheart, 2005). On the other hand, when a word is unfamiliar, reading occurs via the sub-lexical route. In these cases, reading is slower and requires careful phonetic decoding to "hear" the word, thus accessing the semantic lexicon and meaning through the phonological buffer and output lexicon. Sub-lexical reading describes the method by which children first learn to read, and it is also how expert readers approach unfamiliar words (Awaida & Beech, 1995). The lexical route starts with the orthographic input lexicon, which holds entries for words whose written form is known to the reader. In this lexicon, words are organized by frequency. The information then flows to the phonological output lexicon, where the reader accesses the phonological form of the word. The last stage of the lexical route, which is shared with the sublexical route, is the phonological output buffer, where stems and affixes are recombined. From the orthographic input lexicon, the lexical route is also connected to the semantic lexicon and the conceptual system, which allows for comprehension (Friedmann & Coltheart, 2018). Deficits in either of these processing routes can result in specific subtypes of dyslexia.

In cases where the lexical route is impaired, readers overly depend on the sub-lexical route for reading even familiar words, leading to surface dyslexia. Surface dyslexia arises from impairments in the components or connections of the lexical route (Castles, 1996; Coltheart et al., 1983; Friedmann & Lukov, 2008, 2011; Zoccolotti et al., 1999). Readers with surface dyslexia struggle to read words whose pronunciation cannot be easily predicted by grapheme-phoneme conversion rules, such as irregular words and words with multiple phonological options. They often regularize irregular words (e.g., pronouncing the */t/* in *listen*), particularly when reading sub-lexically creates another real word (e.g., reading *bear* as *beer*). Reading pseudowords and regular words is unaffected in pure surface dyslexia (Güven & Friedmann, 2022).

A deficit in the sub-lexical route can lead to grapheme-phoneme conversion dyslexia (sometimes called phonological dyslexia in the literature). Grapheme-phoneme conversion dyslexia involves difficulties in converting graphemes into phonemes (Beauvois & Derouesne, 1979; Campbell & Butterworth, 1985; Castles & Coltheart, 1993). Readers with this dyslexia subtype struggle to decode pseudowords or unfamiliar words, relying heavily on their mental lexicon rather than phonetic decoding (Castles & Coltheart, 1993). Vowel dyslexia, a more specific subtype deficit in sub-lexical reading, is characterized by errors involving the omission, migration, substitution, or addition of vowel letters, while consonants are typically read correctly (Traficante et al., 2021). Vowel dyslexia primarily affects pseudoword reading but can also impact real words when co-occurring with surface dyslexia, as readers must rely on the sub-lexical route (Güven & Friedmann, 2021; Khentov-Kraus & Friedmann, 2018).

Damage to both the lexical and sublexical reading routes can lead to deep dyslexia, forcing readers to access written words predominantly through semantic mediation rather than through reliable orthographic-phonological mappings (Ellis & Young, 1996; Friedmann & Coltheart, 2018; Friedmann & Haddad-Hanna, 2014). Under these circumstances, the activation reaching the semantic system from print is underspecified, increasing competition among semantically related candidates and giving rise to errors such as reading *sand* as "beach". This does not imply that the semantic system itself is damaged – as evidenced by the absence of comparable semantic errors in spontaneous speech or picture naming – but rather that the input arriving from degraded lexical and sublexical pathways is insufficiently constrained. Consequently, imageability becomes critical: concrete, highly imageable words are read more accurately than abstract or function words, and morphologically complex words often lose or substitute affixes. Pseudowords, which cannot be resolved through meaning, are typically rejected or lexicalized (e.g., reading *diger* as "tiger"). In this view, the hallmark semantic errors of deep dyslexia reflect the challenges of accessing meaning indirectly and imprecisely when both primary reading routes are compromised, rather than a breakdown of the semantic system itself.

At the convergence of the lexical and sublexical reading pathways, the phonological output buffer functions as a short-term memory store responsible for holding and assembling phonological word units. This buffer is essential for both lexical and sublexical reading processes (Dotan & Friedmann, 2015; Guggenheim & Friedmann, 2014). Individuals with phonological output buffer dyslexia exhibit a deficit in this system, resulting in difficulties when reading aloud longer words, morphologically complex forms, or pseudowords (e.g., *dancing*, *binking*). Within the phonological buffer, certain items – such as function words, numbers, and some morphemes – appear to be represented as whole units. This may explain their frequent substitution with other items from the same grammatical or categorical class (Cohen Patrick Verstichel Stanislas Dehaene, 1997; Dotan & Friedmann, 2015). Notably, these impairments are typically limited to tasks that require spoken output, such as reading aloud or other purely expressive language tasks. Such tasks are therefore essential for assessing this form of dyslexia, especially when the goal is to distinguish function word or morphological errors from deficits in semantic processing – such as those seen in deep dyslexia, including function word substitutions or derivational errors (e.g., reading *beauty* as *beautiful)*.

Popular methods for dyslexia screening

Research on multiple dyslexias (note the plural in "dyslexias") has shown that different forms of dyslexia can occur independently or coexist within the same individual (Brunsdon et al., 2002; Friedmann & Coltheart, 2018; Sprenger-Charolles et al., 2000; Zoubrinetzky et al., 2014). Despite this evidence, many commonly used screening tools in both research and clinical settings tend to filter for a narrow definition of dyslexia. As a result, these tools may fail to detect certain dyslexia subtypes, increasing the risk of false negatives (Scarborough, 1998; Torgesen, 2002). For instance, the phonological deficit theory emphasizes that dyslexia stems from difficulties in phonological development – specifically, the ability to identify, store, retrieve, or manipulate syllables and phonemes (Goswami, 2011; Ramus et al., 2003; Snowling & Melby-Lervåg, 2016; Tallal, 1980; Ziegler et al., 2008). Screeners based on this theory typically rely on oral-based tasks like rapid automatized naming (RAN), phoneme manipulation, and nonword repetition. However, these assessments may overlook individuals with dyslexia who demonstrate intact phonological processing (Friedmann et al., 2010, 2012; Güven & Friedmann, 2019; Kezilas et al., 2014; Kohnen et al., 2012; Potier Watkins et al., 2023; Traficante et al., 2021).

A narrow focus in screening also risks lumping distinct dyslexia subtypes together, which can lead to inappropriate interventions. For instance, screeners focused on deficient visuospatial attention assess these deficits using tasks that involve rapid serial presentation, where letters or symbols are flashed quickly, and participants must recall as many items as possible (Bosse et al., 2007; Valdois et al., 2004; Vidyasagar & Pammer, 2010). These tasks claim to detect visual-attention-related difficulties that affect letter identification, letter position coding, and the ability to process letter strings in parallel. However, this approach overlooks the evidence that errors in letter identification and letter migrations within or between words can occur independently (Potier Watkins et al., 2023). This is problematic because interventions for different dyslexia subtypes, such as using a finger to follow along for letter position dyslexia or employing a window cut-out to separate words for attentional dyslexia, are specific to each subtype (Friedmann & Rahamim, 2014). Failing to differentiate between subtypes may lead to inappropriate or ineffective interventions.

Encouragingly, an increasing number of screeners used in clinical settings now recognize the multiple dyslexias hypothesis and have expanded to assess both lexical and sub-lexical reading processes, as described by the Dual Route Model. These screeners typically involve reading separate word lists to target these two pathways: surface dyslexia (a deficit in the lexical route) is assessed through irregular word reading, while grapheme-phoneme conversion dyslexia (a deficit in the sub-lexical route) is evaluated using pseudoword reading tasks. Examples of widely used screeners include BALE (Jacquier-Roux et al., 2010), BELEC (Mousty et al., 1994), Exalang (Thibault et al., 2003) and Evaleo (Launay et al., 2018) in French, and DST-J (Fawcett & Nicolson, 2004) and KTEA-3 (Kaufman & Kaufman, 2014) in English. These screeners rely on accuracy and reading speed for the different lists, limiting their ability to provide detailed insights into specific dyslexia subtypes. For instance, a deficit in pseudoword reading, often attributed to grapheme-phoneme conversion dyslexia, could actually be due to earlier visual processing difficulties. In such cases, interventions focused solely on grapheme-phoneme training may be ineffective, while strategies such as increased letter spacing or using a word window to isolate words during reading would be more beneficial (Alexander et al., 1991; Friedmann & Rahamim, 2014). Additionally, these screeners risk producing false negatives by failing to include stimuli that target the specific patterns of errors characteristic of each dyslexia subtype. While the expansion of assessment methods marks progress, the focus on accuracy and speed can still lead to missed diagnoses or the incorrect grouping of different dyslexia subtypes under a single category. To effectively test for multiple dyslexias, it is necessary to provide a broader range of stimuli that address various impairments, including those related to peripheral visual processing. Moreover, scoring should account for the exact types of errors made by readers and not only the percentage of errors on specific items, to better identify the underlying causes of their difficulties.

Designing screeners for clinicians to assess for multiple subtypes of dyslexia

There are, however, a small number of batteries based on the study of the exact errors made by readers, which allow for the fine-grained identification of different subtypes of dyslexia: the Tiltan battery in

Hebrew (Friedmann & Gvion, 2003) used in research, or the Motif battery in English (McArthur et al., 2023) which can also be used by teachers and clinicians. In French, Potier Watkins and Dehaene developed the Malabi dyslexias screener. Using the Malabi, they identified and reported the first French cases of selective attentional dyslexia and letter-position dyslexia (Potier Watkins et al., 2023). The design of the Malabi was inspired by the Dual Route Model of reading, the evidence for selective deficits in the literature, and the Tiltan reading battery in Hebrew (Friedmann & Gvion, 2003). The test includes three reading lists: single word reading, pseudoword reading, and word pair reading. Each word in the Malabi was carefully selected to identify specific subtypes of dyslexia, and the readers' errors were systematically recorded. The test required participants to read a total of 289 words. To identify a selective dyslexia, the authors counted errors that were made in a significantly greater number from normal reading for a particular subtype using a Crawford and Garthwaite's t-test for the comparison of an individual to a control group (Crawford & Garthwaite, 2007).

However, certain limitations prevent the Malabi from being used in a clinical setting. Its length could pose a significant challenge for dyslexic readers, particularly younger children, whose errors on the long list might stem from fatigue rather than dyslexia. Additionally, while the Malabi included items designed to elicit specific types of errors, it did not control for the frequency of grapheme-phoneme correspondences (GPCs). This lack of control meant that many items featured more complex GPCs, which, for younger readers, could be more reflective of inexperience rather than dyslexia. It should be noted that this research highlighting selective cases of dyslexia in the Malabi was recorded in middle-school students (age 11 to 13 years old) (Potier Watkins et al., 2023).

Additionally, most real-world reading, whether in school or daily life, involves continuous text rather than isolated word lists. In continuous text reading, eye movements play a critical role in fluency (Hindmarsh et al., 2021; Rayner & McConkie, 1976). Dyslexic readers often exhibit longer fixation times, more frequent fixations, shorter saccades, and skipped words, indicating a different approach to text processing compared to typical readers (Benfatto et al., 2016).

To address these limitations, we developed the Mariette screener – evaluated in this paper – designed to comprehensively track both reading progress and specific deficits throughout elementary school education. The Mariette was written in a nonsense text format, preventing reliance on context for word recognition. This means the text follows correct syntactic structures but uses out-of-context words, pseudowords, or rare words (e.g., an example sentence, translated to English: *Near the truche, a child breaks pebbles into ten reliable squares*). Nonsense texts require the child to read from left to right and using similar eye movements as in text reading while ensuring that words cannot be guessed based on context (Friedmann & Rahamim, 2007; Lefavrais, 2005).

The Mariette was designed with a total of 294 items, 16 of which were pseudowords, and included a 5-minute stop-time which serves as a cutoff for cases of severe dyslexia or with young children who are still learning to read. This length was chosen to closely match the Malabi screener for older readers while also aligning with the popular French nonsense dyslexia screener, the Alouette, which contains 275 words and uses a time-stop approach to screen for dyslexia (Lefavrais, 2005). The Mariette's design allows it to be used consistently across various stages of reading development, providing standards for younger students who read fewer words while still using the same test. As students progress through the Mariette, the complexity of grapheme-phoneme correspondences (GPCs) increases, and less common irregular words are introduced. According to France's National Evaluation guidelines, by the end of 1st grade (start of 2nd grade), students are expected to be familiar with basic GPCs and common irregular words, with most first graders able to read at least 75 words in three minutes in a nonsense text (Lefavrais, 2005). To ensure the content is appropriate for each grade level while still detecting selective dyslexic errors, the first paragraph of the Mariette comprises 75 words featuring GPCs and irregular words suitable for grade 1 reading materials (Lété et al., 2004). As students advance to higher grades, their reading speed increases, allowing them to read more words within the same time limit. Therefore, subsequent paragraphs introduce more complex GPCs and rarer irregular words, gradually increasing the difficulty while incorporating a larger number of dyslexia-sensitive items. This design captures various forms of dyslexia while providing age-appropriate content that aligns with students' reading development.

To design the Mariette, we focused on some of the subtypes of dyslexia described in the previous paragraphs, for which there is sufficient literature to justify the choice of stimuli. We selected words and pseudowords that, when misread due to a specific subtype of dyslexia, transform into another valid word that, furthermore, still fits syntactically within the sentence. For example, in the case of letter-position dyslexia, the verb *rogné* in the Mariette (*"trimmed"* in English) may be misread as *rongé* (*"gnawed"*) if the letters "*g*" and "*n*" are transposed. This approach increases the screener's sensitivity to detect specific dyslexia subtypes, as readers are less likely to self-correct if they read a known word. Table 1 provides an overview of the different types of errors, associated with different subtypes of dyslexia, that can be identified using the Mariette and examples of stimuli specifically selected to target each type of error.

Using the Mariette, we conducted two studies presented in this paper. The first study explores the normal developmental trajectory of reading errors, drawing from the literature on developmental dyslexia. The second study evaluates the Mariette's effectiveness as a diagnostic tool in clinical settings.

Study 1: evolution of different type of errors in normal readers over the course of schooling

Methods

Participants

To characterize normal reading development using the Mariette and establish norms, 812 children from 1st to 5th grade (age: 5 years 11 months to 11 years 11 months) were recruited from 7 schools of diverse backgrounds, for a representative sample: 38% of the children tested lived in rural areas and 27% of the children came from low socio-economic status schools, called priority education schools in France. These percentages were slightly higher than at the national level, where 20% of children live in rural areas (Direction de l'Evaluation, de la Prospective et de la Performance, 2021) and 20% of children are in priority education schools (Stefanou, 2022). Only typically developing children, based on teachers' reports, were included. 1.8% of children were excluded because their age was greater than or less than two standard deviations from the mean, resulting in a subsample of 797 children from 1st to 5th grade (age: 5 years 11 months to 10 years 2 months).

Materials

The Mariette. The Mariette screener consists of a standardized nonsense text comprising 294 items, presented on A4 paper in Calibri 14-point font with 1.5 line spacing. Students are allotted five minutes to read as much of the text as possible. As outlined in our introduction, the screener is structured into four paragraphs of progressively increasing difficulty, specifically calibrated to challenge grapheme-phoneme correspondences (GPCs) and irregular word recognition. Each paragraph contains carefully selected items designed to identify the specific dyslexia subtypes detailed in Table 1. To ensure controlled assessment conditions, we maintained consistency across paragraphs in factors that should remain stable across grade levels: all pseudowords were constructed to be phonologically decodable; irregular words were characterized as words for which the application of French grapho-phonemic conversion rules results in a pronunciation different from the actual pronunciation (for example, reading the irregular word "temps" (/tɑ̃/) as /tɑ̃p/); and paragraphs were matched for word length (Kruskal–Wallis test, $p = 0.45$) and word frequency (Kruskal–Wallis test, $p = 0.62$), with the deliberate exception of irregular words. We deliberately included frequent irregular words in the first paragraphs – those read even by the youngest children – to ensure that any regularization of irregular words could not be explained solely by limited vocabulary. This design allows for targeted assessment of reading development while isolating potential deficits associated with specific dyslexia subtypes.

The Alouette. To assess the reliability of this test, 10% of randomly sampled students (90 students, 15 1st graders, 19 2nd graders, 15 3rd graders, 24 4th graders and 17 5th graders) also took the Alouette test (Lefavrais, 2005). The Alouette text was presented in the same format as the Mariette, typed in Calibri 14-point font with 1.5 line spacing. The images in the original text were not reproduced. The Mariette screener was modelled after the Alouette test, a popular dyslexia assessment tool in France. We selected this established instrument for its comparable design featuring meaningless text of similar word count (265/294), which effectively evaluates decoding skills without contextual support. Building on this proven approach, the Mariette extends diagnostic capabilities to identify specific dyslexia subtypes. Students had 3 min to read as many words as possible in the Alouette test. To compare the Mariette and Alouette tests across all grades, we computed reading speed as the number of words correctly read per minute and accuracy, as the percentage of correctly read words.

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 child by the teacher. Participants were tested in a quiet room by our research team at school during school time. The child was told: *I'm going to show you a story that might seem a bit strange, so it's completely normal if parts of it don't make sense. There will also be some made-up words, so don't worry if you're unfamiliar with them. I'd like you to read the story aloud, doing your best, and see how far you can get in 5 min. I'll start the timer. Ready? Let's*

Table 1. The different types of error recorded in the Mariette.

Error type	Description	Stimuli sensitive to that type of error	Example → possible error
Attentional	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 (9 pairs in the text) 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).	*Addition* Page dira → page dirage *Omission* plante étrange → plate étrange *Substitution* lame rime → rame rime
Letter position	Letter transposition within words and pseudowords. Omission or substitution of an instance of a repeated letter, or repeating of a letter	All items in the text, and in particular transposable words and pseudowords (20 in the text): items for which a transposition in the word forms a word.	*Transposition* rogné → rongé arbi → abri *Repeating a letter* sa**r**dine → sa**r**d**r**ine *Omission of a repeated letter* p**r**i**r**ent → pi**r**e *Substitution of a repeated letter* pa**g**e → **g**a**g**e
Other visual errors	Omissions, substitutions and additions of letters resulting neither from a letter position error nor from an attentional error.	All text items	*Addition* mare → marche *Omission* compris → copris *Substitution* lange → lande
Voicing	Pronunciation of a voiceless consonant as its voiced counterpart (e.g., /t/ as /d/, /f/ as /v/, /k/ as /g/), and vice versa.	All text items	dents → tant
Vowel	Subexical deficit only affecting vowel reading. Errors are substitutions, additions, omissions and transpositions of vowels only.	All text items	flache → flèche
Regularization	Regularization of letter sound in irregular words resulting from a failure of the lexical process in which the sublexical process produces high probability grapho-phonemic mappings instead of the (lower probability) lexically correct mapping.	Irregular words (22 in the text)	Reading “temps” (/tɑ̃/) as /tɑ̃p/
Contextual rule	Incorrect pronunciations of specific letters (such as s, g, c, or t in French) whose sounds vary depending on the surrounding letters.	Item containing contextual letters (46 in the text)	Reading “dragée” (/draʒe/) as /drage/
Mispronunciation of digraphs	Pronouciation of the sound of each individual letter within a digraph (a two-letter grapheme) separately instead of as a unit	Words containing digraphs (159 in the text)	Reading “rogné” (/roɲe/) as /rogne/
Function word substitutions	Substitution of a function word (closed class words: determiners, pronouns, prepositions and conjunctions) with another one from the same grammatical category.	Function words (123 in the text)	le → les, la, du
Morphological errors	Derivational errors (addition of a prefix or suffix creating a new word or changing its grammatical category or meaning) and inflectional errors (omission of conjugation, plural or feminine markers).	All text items except function words and in particular on morphologically complex words (29 in the text)	*Derivation errors* étrange → étrangère *Inflection errors* arrivèrent → arrive

begin!. The researcher turned over the page and started the timer. Each child’s reading was recorded for error analyses.

Given the young age of some children in our sample, we did not stop them from using their finger. No feedback was given, but, if needed, they were encouraged to continue. For those tested at the end of the first trimester and mid-way through 1st grade, half of the text was concealed to avoid discouraging their efforts, as reading the full text was beyond their expected grade level. For the 10% of randomly selected children who also took the Alouette test, it was always administered after the Mariette. Instructions closely resembled those used for the Mariette.

In France, 1st and 2nd grade are key years for formal reading instruction, with a strong focus on learning the alphabetic code of the language. For this reason, children in these grades were tested at three points in the year to capture their progress during these critical periods of learning: at the end of the first trimester (December 2022), mid-year (March 2023), and at the end of the school year (June 2023). All other grades were only tested in the middle of the school year (March 2023). Testing took place over a period of 2 to 4 weeks. The exact number of children tested for each test session is shown in Table 2.

Error type scoring. Reading errors were transcribed phonetically. When the child made multiple responses to all or part of the stimulus, the first response to each syllable was transcribed. Occasionally, children 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 child. To characterize error types, we then applied a procedure described by Potier Watkins et al. for the Malabi tests (2023). For each error, a specific error type was assigned by consensus among two researchers involved in the project, and the information was entered into a database. This approach allowed us to partially automate the error classification process – once a particular error was identified and added to the database, it could be automatically applied to similar cases. Below we describe how an error was attributed to a particular error type.

Letter position errors were recorded when a transposition was made that included letters in the word (for example, when the word *fiable* (*"reliable"* in English) was read *faible* (*"weak"*)), when omitting or substituting an instance of a repeated letter (omission: reading *bible* (*"bible"*) as *bile* (*"gall"*) – substitution: reading *page* (*"page"*) as *gage* (*"pledge"*)) or repeating a letter (reading *bile* (*"gall"*) as *bible* (*"bible"*)). An addition, omission or substitution of a letter (or group of letters) in a word was categorized as an attentional error if it involved letters that could have come from a surrounding word (side, top or bottom), and maintained the same relative position from the word of origin (first letter, last letter, inner letter). An addition, omission or substitution that could not be labelled as either attentional or letter position was categorized as an other visual error when it involved a consonant (for example when the word *bois* (*"wood"* in English) was read *mois* (*"month"*)) and as a vowel error when it affected a vowel (for example when the pseudoword *grusée* was read *grisée*). Letter position and attentional errors affecting a vowel were double counted as vowel errors and letter position or attentional errors. Regularization errors were observed during the pronunciation of letter sounds in irregular words (for example, pronouncing the letter "t" in *"listen"*). These errors result from a failure of the lexical process in which the sublexical process produces high probability grapho-phonemic mappings instead of the (lower probability) lexically correct mapping. Contextual rule errors refer to incorrect pronunciations of specific letters (such as s, g, c, or t in French) whose sounds vary depending on the surrounding letters. When the sound of each individual letter within a digraph (a two-letter grapheme) was pronounced separately instead of as a unit (for example, pronouncing both letters in "in" individually rather than as a single nasal sound) we classified this as a digraph error. Regularizations of irregular words, misapplication of contextual rule and digraph errors reflect failures of learning: regularizations reflect failure to learn the lexical representation, while the other two reflect failure to learn complex grapho-phonemic mappings. Thus, in older readers, regularization errors reflect impairments in the lexical route, whereas misapplication of contextual rules or digraph errors arise from failure in the sublexical process. Another error type that may arise from a deficit in sublexical processing is voicing errors, which involve pronouncing a voiceless consonant as its voiced counterpart (e.g., /t/ as /d/, /f/ as /v/, /k/ as /g/), and vice versa. Errors on function words (closed class words: determiners, pronouns, prepositions and conjunctions) were characterized as substitutions within function words when a function word was substituted by another one of the same grammatical category. Finally, morphological errors were classified as either derivational (reading happy as happily) or inflectional (for example, reading *eaten* as *eat*). When an error made by a child was too far removed from the original word to be interpreted (for example, the word *sur* (*"on"*) read as *sorti* (*"out"*)), it was categorized under the label *other*. Using this classification, we were able to characterize 93.9% of the errors encountered, with an increase in this percentage as children get older (96.2% of the 5th graders' errors vs 91.6% of 1st graders' errors).

If an error could reasonably be attributed to more than one category (for example, reading *nous* as *vous*, which could reflect either a function word substitution or a letter substitution (changing "n" to "v")), it was counted under both categories. We refer to these cases as "ambiguous" errors, meaning the exact source of the error is ambiguous, but both categories are relevant. Similarly, when an error results from a sequence of distinct processes (such as reading *pose* as *possé*, which may involve first a contextual rule error followed by a substitution), it was also counted in both relevant categories. We refer to these as "multiple responses" errors, indicating that both error types are clearly

Table 2. Number of children per grade for each measurement period.

Grade	1st	2nd	3rd	4th	5th
End of the first trimester	87	82	/	/	/
Middle of the year	82	88	74	103	104
End of the year	90	87	/	/	/

involved. In both cases, All types of errors were included in the standards, following the principle that a child with a specific subtype of dyslexia will exhibit a higher frequency of certain errors compared to others. When the accumulation of a specific error type exceeds the thresholds set by the standards, it should alert the practitioner to the possibility of a specific dyslexia subtype.

Data analysis. ***Reading speed and accuracy.*** Reading speed and accuracy are two classical parameters measured in text reading tests. Reading speed was calculated as the number of items correctly read per minute and reading accuracy as the percentage of correctly read items out of the total of number of item read. 19 children whose speed or accuracy was more than three standard deviations away from the mean were considered as outliers and removed from the analysis. We studied the evolution of both scores over the course of schooling using linear regressions and Welsh's t-tests.

Self-corrections. The term self-correction refers here to mistakes quickly followed with the correct answer. In some tests, such as the Alouette, self-corrections are considered as correct answers, with the underlying idea that the student's hesitations will still necessarily result in a detectable slowdown in their reading speed. Here however, since we focus on error types, it is essential to keep a record of children's 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 children 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.

Traditional scoring, percentage of errors for irregular and pseudowords. Most current screeners rely on error rates for irregular words and pseudowords to identify surface and phonological dyslexia. To compare the effectiveness of this approach with a more detailed method of categorizing each error made by the reader, we calculated the average accuracy for each child on irregular words (22 words in the text) and pseudowords (16 words in the text). Using the classification described in Table 1, we then analyzed each error on irregular words and pseudowords specifically, determining the proportion of errors that matched the expected type (regularization error for irregular words and grapheme-phoneme conversion errors (contextual rule, digraph and voicing errors) for pseudowords).

Error type scoring, percentage of errors for each type. Because students were stopped after five minutes of reading, they did not read the same number of words. To track how different types of error occurrences changed over time, we calculated each error type as a percentage of the opportunities the student had to make that error type. For example, attentional errors (addition, omission or substitutions involving letters that could have come from a surrounding word) can occur across all words in a text. If a child reads 230 words and makes 15 attentional errors, their attentional error rate would be 6.5%. Conversely, substitutions within function words are limited to function words. A child who encounters 20 function words and substitutes 15 of these would have a rate of substitutions within function words of 75%. A score closer to 100% indicates a consistent error pattern for a specific item type. The items counted for each type of error are described in Table 1. Those percentages were then compared between grades using Welch's t-tests.

Results

Was the Mariette comparable to known reading standards?

We first present the results for the 10% of students who completed both the Mariette and a similar reading screener with established grade-level norms, the Alouette (Lefavrais, 2005). The correlations between the Mariette and the Alouette, graphically depicted in Figure 2, showed that reading speed, measured as the number of words correctly read per minute, was highly correlated across tests ($r = 0.97$, $p<0.001$), as well as reading accuracy ($r = 0.85$, $p<0.001$). These outcomes confirmed that the Mariette measures overall reading performance similarly to the Alouette, while allowing for a much more sophisticated error analysis.

Grade-level development of reading speed and accuracy

Figure 3 depicted how children's reading speed and accuracy changed over the course of schooling. Between the start of 1st grade and the end of 4^{th} grade, reading speed, expressed in words read correctly per minute, increased almost linearly ($r = 0.79$, $p = 0.021$), reaching a plateau between 4^{th} and 5^{th} grade. Pairwise comparisons showed that reading speed increased significantly between each timepoint ($p<0.05$ for each comparison) except between 4^{th} and 5^{th} grade ($t(196.2) = -0.133$, $p = 0.89$).

Accuracy followed a less linear trajectory – it rose sharply during the first year of primary school, increasing from 46.6% to 79.5%, with significant pairwise differences between the beginning, middle, and end of first grade (all $p < 0.001$). This sharp rise likely reflects the transition from having no reading skills to acquiring foundational knowledge through intensive instruction. From the beginning of second grade to fifth grade, accuracy continued to improve, but at a slower pace ($r = 0.92$, $p = 0.009$), eventually reaching 93.9%. This pattern highlights the dramatic gains that occur when children first learn to read – moving from non-reader to reader – compared to the more gradual and variable improvements that follow, which likely reflect fluctuations in mastery across different children and reading demands.

Evolution of self-corrections over the course of schooling

As shown in Figure 4, the percentage of self-corrected errors was low at the start of 1st grade, with only 3% of errors being self-corrected. This percentage increased over the course of 1st grade, and even more sharply between the end of 1st grade and the beginning of 2nd grade, when it rises from 8% to 16%. It then fell during 2nd grade, before rising again in 3rd grade and stabilizing at around 20% between 4th and 5th grade. This increase in self-corrections likely reflects growing metalinguistic awareness and reading skill, as children become better able to recognize and repair their own mistakes – particularly in the unusual task of reading pseudowords, which younger readers often fail to monitor effectively.

Evolution of different error types over the course of learning

Traditional scoring, percentage of errors for irregular and pseudowords. Most standard screeners currently used to identify different subtypes of dyslexia score error rates on irregular words and pseudowords to detect surface and phonological dyslexia. These

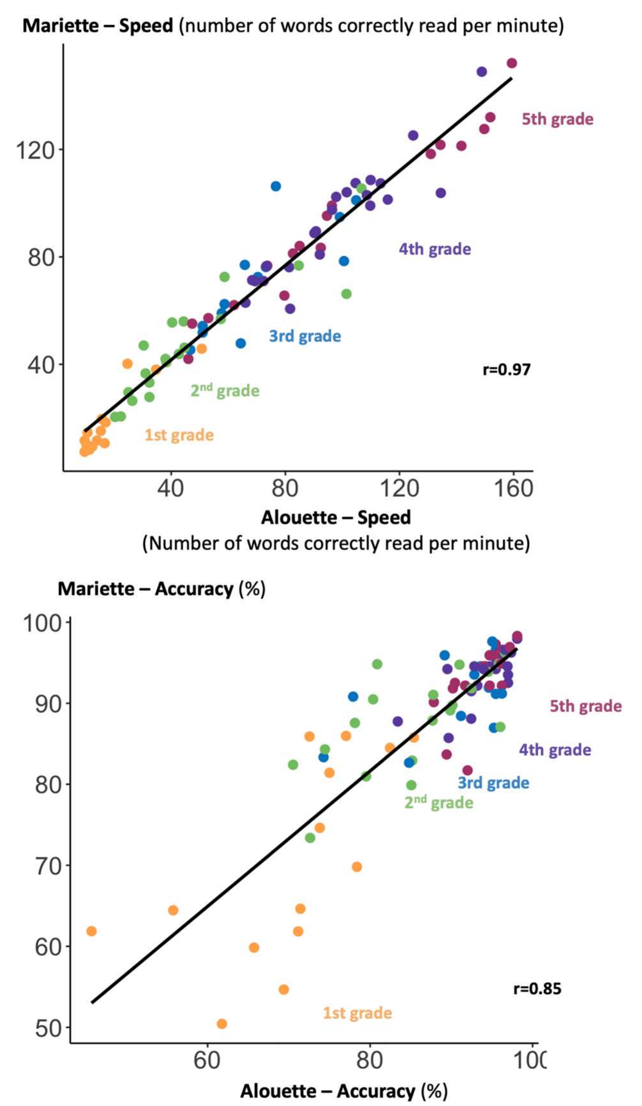

Figure 2. Reading speed and accuracy correlations between the Mariette and the Alouette, a widely used dyslexia screening tool.

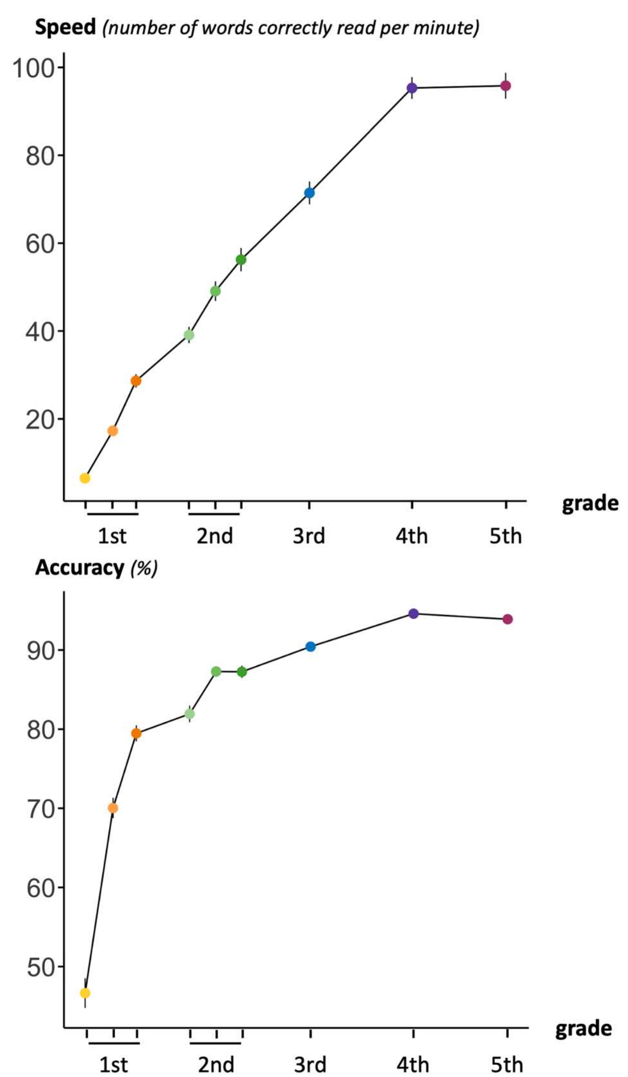

Figure 3. Changes in reading speed and accuracy over the course of schooling. Errors bars represent the standard error of the mean.

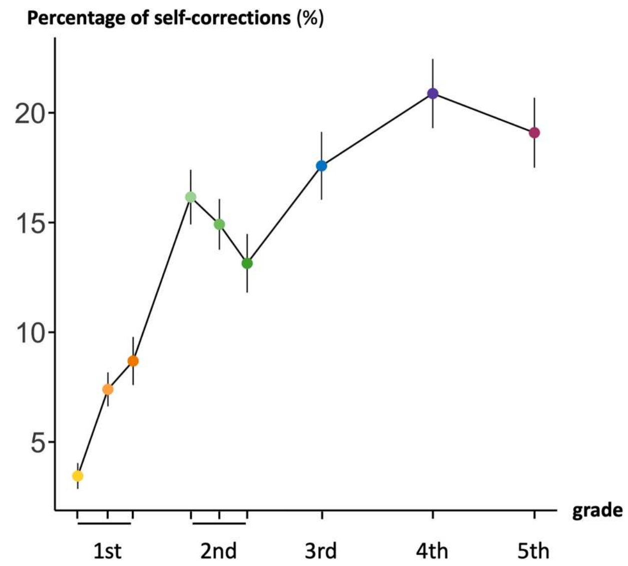

Figure 4. Evolution of the percentage of self-corrections over the course of schooling.

screeners operate on the assumption that errors on irregular words are surface errors, while errors on pseudowords result from grapheme-phoneme conversion issues, per the dual-route reading model. However, our observations showed that, by the middle of first grade, only 57.6% of errors on irregular words could be categorized as regularizations, and just 43.2% of errors on pseudowords were due to grapheme-phoneme conversion difficulties (voicing errors, digraph errors or contextual rule error). In first grade, pseudoword errors were largely related to letter-position errors (27%), while irregular words produced a notable amount of other visual errors (14.1%). Although the proportion of regularization errors on irregular words increased over time, it never reached 100%. By fifth grade, regularization errors still only accounted for 70.4% of errors on irregular words, with visual errors (including letter-position, attentional, and other visual errors) making up 25.7%. Conversely, the percentage of grapheme-phoneme conversion errors among pseudoword errors declined over time, dropping to 12.6% by 5th grade. At this stage, most errors on pseudowords (56.2%) were due to letter-position issues. These findings indicate that the type of word cannot reliably predict the type of error, even in older children.

Error type scoring, percentage of error for each error type. Characterizing the errors allowed us to examine how the frequency of each error type changed over the course of primary school. As shown in Figure 5, the overall percentage of errors decreased from the start of 1st grade to 5th grade, though the rate and pattern of this decline differed by error type.

Other visual errors (additions, omissions and substitutions of letters resulting neither from a letter position error nor from an attentional error) declined sharply during the first year of primary school, with significant decreases between the start and the middle of 1st grade (t(132.7) = 5.25, *p*<0.001), and between the middle and the end of 1st grade (t(148.57) = 4.00, *p*<0.001), followed by a second decline between the start and the middle of 2nd grade (t(123.23) = 2.84, *p* = 0.0052). Subsequently, such errors continued to decline from the end of 2nd grade to 5th grade, with significant drops occurring between the end of 2nd grade and 3rd grade (t(155.18) = 2.31, *p* = 0.022), and between 3rd and 4th grade (t(105.72) = 4.14, *p*<0.001). Attentional errors (addition, omission or substitutions involving letters that could have come from a surrounding word) followed a similar pattern to other visual errors, with a sharp decline in first grade (start of 1st grade - middle of 1st grade: t(161.96) = 3.18, *p* = 0.0018; middle of 1st grade - end of 1st grade: t(127.72) = 4.20, *p*<0.001), followed by a second decline between the start and middle of 2nd grade (t(134.54) = 3.21, *p* = 0.0017), between the end of 2nd grade and 3rd grade (t(146.05) = 2.79, *p* = 0.0060) and between 3rd and 4th grade (t(128.58) = 3.51, *p*<0.001).

Letter position errors (letter transposition within the word, omission or substitution of an instance of a repeated letter or repeating of a letter) followed a slightly different pattern, with a significant decrease in 1st grade (start of 1st grade - middle of 1st grade: t(133.76) = 3.30, *p* = 0.0012; middle of 1st grade - end of 1st grade: t(135.66) = 3.80, *p*<0.001). This was followed by a slower decline, with the only further significant reduction between 3rd and 4th grade (t(148.35) = 2.81, *p* = 0.0057) and levelling-off between 4th and 5th grades.

Interestingly, we observed a trend where young readers exhibited a higher percentage of attentional errors compared to letter position errors, though this difference was not statistically significant (start of 1st grade: t(151.84) = 0.45, *p* = 0.66; middle of 1st grade: t(153.51) = 1.28, *p* = 0.20; end of 1st grade: t(172.95) = 0.77, *p* = 0.44). This trend reversed in 2nd grade, where letter position errors then became significantly more frequent than attentional errors (start of 2nd grade: t(155.97) = −2.16, *p* = 0.033; middle of 2nd grade: t(160.05) = −4.62, *p*<0.001; end of 2nd grade: t(167.45) = −2.95, *p* = 0.0037; 3rd grade: t(130.24) = −4.65, *p*<0.001; 4th grade: t(159.28) = −5.58, *p*<0.001; 5th grade: t(154.04) = −7.62, *p*<0.001).

Looking at central reading processes, students at the start of 1st grade made a lot of regularizations of

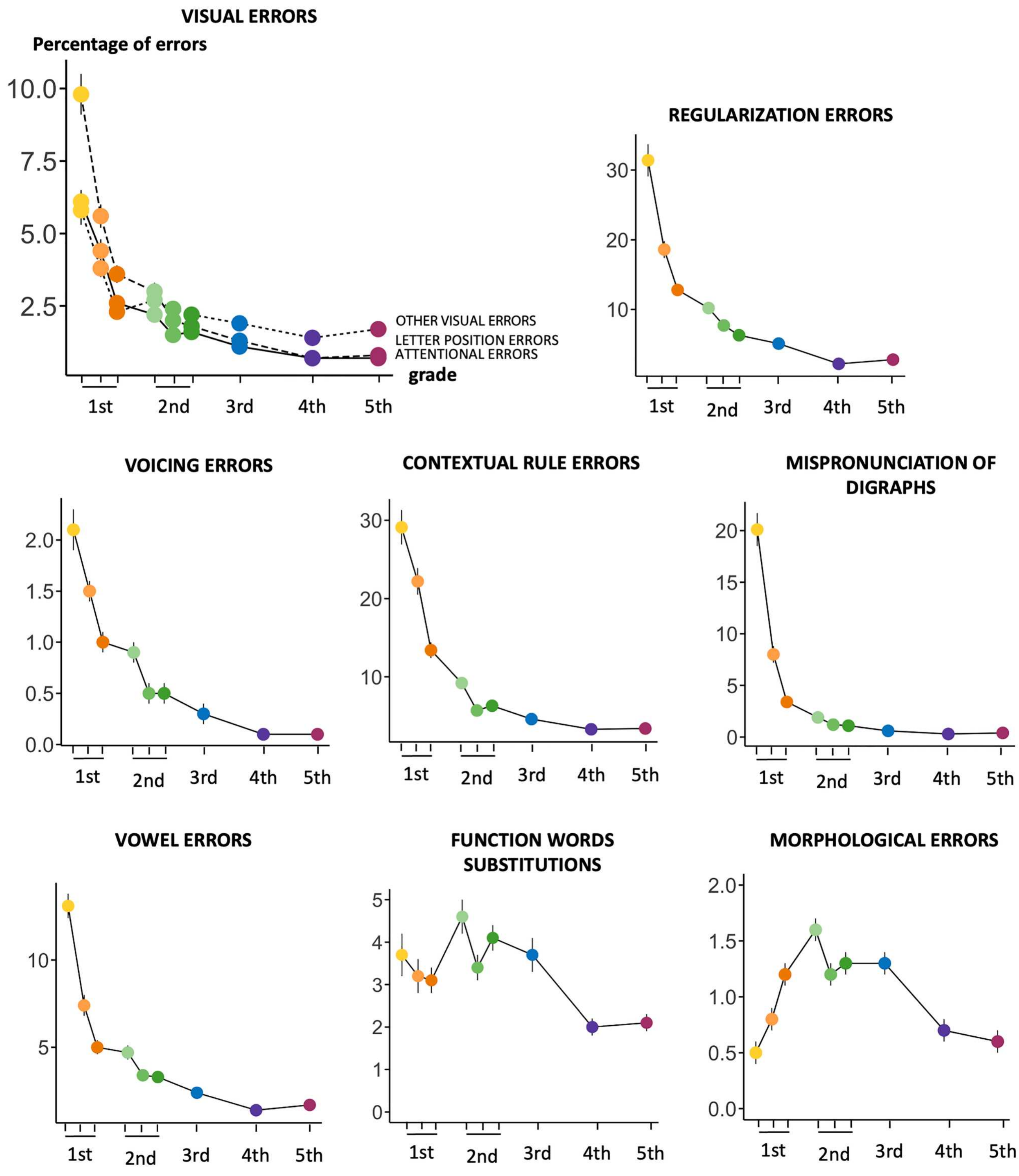

Figure 5. Evolution of different types of error over the course of schooling. Percentages were calculated in proportion to the number of possible occurrence of the errors (instead of the total number of errors). For example, 100% of regularization errors means that each time the child encountered an irregular word, he regularized it. This is why the sum of the percentages in the graphs above does not equal 100.

irregular words (31.4% on average), digraphs errors (20.1% on average) and contextual rules errors (29.1% in average). These three types of error follow similar trajectories. They fall sharply from the start of 1st grade to the middle of 2nd grade, when they reach 7.7%, 1.2% and 5.7% respectively on average (*regularizations*: start of 1st grade – middle of 1st grade: $t(125.64) = 4.96$, $p<0.001$; middle of 1st grade – end of 1st grade: $t(131.82) =$

4.11, $p<0.001$; end of 1st grade – start of 2nd grade: t(149.53) = 2.95, $p = 0.0037$; start of 2nd grade – middle of 2nd grade: t(157.48) = 3.95, $p<0.001$ / *digraph errors*: start of 1st grade – middle of 1st grade: t(122.51) = 6.63, $p<0.001$; middle of 1st grade – end of 1st grade: t(100.04) = 5.13, $p<0.001$; end of 1st grade – start of 2nd grade: t(155.35) = 3.97, $p<0.001$; start of 2nd grade – middle of 2nd grade: t(141.18) = 2.93, $p = 0.0039$ / *contextual rule errors:* start of 1st grade – middle of 1st grade: t(151.7) = 2.46, $p = 0.015$; middle of 1st grade – end of 1st grade: t(129.99) = 4.52, $p<0.001$; end of 1st grade – start of 2nd grade: t(143.32) = 3.63, $p<0.001$; start of 2nd grade – middle of 2nd grade: t(143.58) = 4.65, $p<0.001$). As with the other error types described earlier, the percentage of these three type of errors continued to decline, but at a slower pace, with a notable drop occurring between the end of 2nd grade and 3rd grade (*regularizations*: t(152.94) = 2.26, $p = 0.025$; *digraph errors*: t(118.92) = 2.84, $p = 0.0053$; *contextual rule errors*: t(148.04) = 3.13, $p = 0.0021$) and between 3rd and 4th grade (*regularizations*: t(112.97) = 6.94, $p<0.001$; *digraph errors*: t(122.08) = 2.18, $p = 0.031$; *contextual rule errors*: t(140.93) = 3.20, $p = 0.0017$). 5th graders still made 2.8% of regularizations, 3.4% of contextual rule errors but only 0.4% of digraph errors on average in the whole text.

Voicing errors had one of the lowest frequency among the error types observed. In the start of 1st grade, children made an average of 2.1% of errors of this type, but this percentage decreased sharply to 0.1% by 5th grade. The decline was statistically significant between the start and middle of 1st grade (t(132.37) = 2.12, $p = 0.036$), between the middle and the end of 1st grade (t(156.69) = 2.56, $p = 0.011$) and continued in later years, with further significant drops between the start and middle of 2nd grade (t(125.29) = 2.94, $p = 0.0039$) and between 3rd and 4th grade (t(94.29) = 3.04, $p = 0.0030$).

We also looked at vowel errors (including additions, omissions, substitutions or transposition of vowels). Vowel errors followed a similar trajectory as grapheme-phoneme conversion errors, with a significant drop from the start to the end of 1st grade (start of 1st grade – middle of 1st grade: t(154.48) = 6.23, $p < 0.001$; middle of 1st grade – end of 1st grade: t(134.87) = 3.58, $p<0.001$), the start of 2nd grade to the middle of 2nd grade (t(131.01) = 2.88, $p = 0.0047$), and again between the end of 2nd grade and 3rd grade (t(153.05) = 2.90, $p = 0.0042$) and between 3rd and 4th grade (t(121.81) = 4.00, $p<0.001$). Morphological errors and substitutions of function words followed a more complex, non-linear pattern compared to other error types. Morphological errors – including both inflectional and derivational changes – increased significantly between the end of 1st grade and the start of 2nd grade (t(164.7) = −2.27, $p = 0.025$), then declined between the start and middle of 2nd grade (t(154.58) = 2.17, $p = 0.031$), and again from 3rd to 4th grade (t(114.71) = 3.44, $p < 0.001$), reaching just 0.5% by 5th grade. Function word substitutions showed a similarly irregular trajectory: they remained stable during 1st grade, increased between the end of 1st grade and the start of 2nd grade (t(150.42) = −2.97, $p = 0.0034$), decreased from the start to the middle of 2nd grade (t(150.95) = 2.49, $p = 0.014$), and declined further from 3rd to 4th grade (t(112.66) = 4.13, $p < 0.001$). However, these errors remained relatively frequent, persisting at around 2% in 5th grade.

Discussion

In this first study, we introduced the Mariette, a screening tool designed to target specific error types while measuring both reading speed and accuracy. We demonstrated that the number of words read correctly per minute using the Mariette was highly correlated with results from a screener frequently used by French speech therapists, the Alouette. The error analysis provided by the Mariette also allowed for a deeper investigation into how typical reading errors, commonly associated with dyslexia, evolve as children progress through school.

The changes in reading speed observed between grades in our results were consistent with other longitudinal and cross-sectional studies, showing a near-linear increase during the early grades (1st and 2nd grade), followed by a plateau towards the end of primary school (Al Otaiba et al., 2009; Biemiller, 1977; Dehaene-Lambertz et al., 2018; Lervag & Hulme, 2009; Nese et al., 2013). Research has repeatedly shown that reading speed, or fluency, is a strong indicator of a child's reading ability and is highly correlated with reading comprehension (Fuchs et al., 2001; Hudson et al., 2005; Lee & Chen, 2019) and with the development of the brain's reading networks (Dehaene et al., 2010, 2015; Dehaene-Lambertz et al., 2018; Feng et al., 2022). On the Mariette, reading speed increased steadily across grades, with significant growth at each testing period, except between 4th and 5th grade. The growth observed between the end of 1st grade and the beginning of 2nd grade was smaller than that seen within the school year, likely due to the absence of explicit instruction during the summer vacation. Nevertheless, the significant progress made between the end of 1st grade and the beginning of 2nd grade suggests that children benefit from "self-learning" during breaks (Share, 1995).

Reading accuracy – another key metric closely linked to comprehension – showed a similar developmental trajectory, particularly in opaque orthographies like French, which typically require a longer period to master (Ziegler et al., 2010). In such languages, it often takes several months for accuracy rates to approach 100% (Seymour et al., 2003; Vaessen et al., 2010; Ziegler & Goswami, 2005). By 4th grade, children reached a plateau of around 95% accuracy, reflecting the automaticity of reading typically achieved by the end of primary school. However, about 5% of errors persisted, likely due to the presence of rare irregular words (e.g., "*dolmen*") and pseudowords resembling common French words (e.g., "*truche*", "*arbi*", "*oviles*", "*mantou*"). The increase in self-corrections from 1st to 5th grade may also be explained by instances where pseudowords are read as real words, particularly when a letter transposition in the pseudoword forms a familiar word. This effect is likely driven by increased reading automaticity and the lexicality effect, whereby pseudowords that closely resemble familiar words are more likely to trigger the retrieval of real word forms from memory (Acha & Perea, 2008).

Most standard screeners currently used to identify subtypes of dyslexia compare error rates on irregular words and pseudowords. Surface dyslexia is identified when the error rate on irregular words is abnormally high, while performance on pseudowords remains within the normal range. Phonological dyslexia shows the opposite pattern, with elevated errors on pseudowords but not on irregular words. Mixed dyslexia is diagnosed when both error rates exceed normative thresholds (Castles & Coltheart, 1993; Manis et al., 1996; Sprenger-Charolles et al., 2000, 2011; Stanovich et al., 1997; Ziegler et al., 2008). These screeners assume that errors on irregular words would be surface errors and errors on pseudowords should be grapheme-phoneme conversion errors. This however, was not what we observed when looking at the types of errors readers make: across all grades, 37% of errors on irregular words and 74% of errors on pseudowords were not of the expected error type. Thus, relying solely on error rates to compare performance across word categories provides a limited picture. To better capture the nature of children's difficulties, it was essential to examine the specific types of errors they made.

In each of the visual error types we assessed (attentional errors, letter-position errors or other visual errors), the percentage of errors improved by grade. There was a significant reduction in these errors between 1st and 2nd grade, followed by another drop between 3rd and 4th grade. Our findings on letter-position errors align with Paterson et al., who also found a decline in error rates between 2nd-4th graders and adults (Paterson et al., 2015). However, unlike Grainger et al. (2012), letter-position errors did not follow a bell curve with a peak in 3rd grade, likely because we did not include enough items specifically focusing on transposable pseudowords, which was the focus of their study. Despite their reduction, letter position errors were still present among 5th grade students, a result that could be partly explained by the transposable items contained in the text (such as "*arbi*", "*oviles*", or "*fiable*"), which, incidentally, are the items on which the majority of letter position errors occur among older students (in 1st graders, 29.8% of letter position errors were made on these items, a percentage that increases with schooling to reach 52.4% in 5th grade). We also observed that the burden of error was not the same for attentional and letter transposition errors. In a crossover effect midway through 2nd grade, errors that could be scored as attentional errors were slightly greater than letter position errors, which were significantly more frequent after this period. This progression in visual error types reveals an interesting developmental path: attentional errors are notably more frequent at the outset, suggesting that early reading may place a heavier burden on general visual attention. However, after the first year of learning, letter-position errors appear to be more challenging, possibly indicating a shift toward finer-grained positional processing. This shift may point to an early peripheral reorganization in how visual attention and positional encoding are applied during reading, reflecting the developmental refinement of focal processing as reading becomes more automatic.

When examining central reading processes, we observed a high percentage of regularizations of irregular words in young children, as to be expected given that at the time of testing, 1st graders were still in the early stages of learning to read and relied heavily on the sub-lexical route to decode words, as their mental lexicon would still be very small (Bijeljac-Babic et al., 2004; Burani et al., 2002; Schröter & Schroeder, 2017; Zoccolotti et al., 2005). At the start of the year, 1st graders also made a lot of digraph errors and contextual rule errors, a pattern that is entirely normal because they didn't learn all grapheme-phoneme correspondances and contextual rules yet. The sharp decline of those type of errors throughout primary school reflects the progressive consolidation of the lexical route, as documented in previous research. However, despite the expectation that 5th graders have mastered reading, we found that they still made an average of 2.8% of regularizations and 3.4% of contextual rule errors while reading the Mariette text. This suggests that the mastery of French contextual rules and irregular word reading continues to pose challenges for older students – consistent with findings showing persistent misapplication of contextual rules even among 6th graders (Lubineau et al., 2024). Our findings also showed that other types of errors, such as voicing errors or digraph errors almost disappeared by the end of 2nd grade. Their continued presence in older students, even in small numbers, may point to residual reading difficulties that warrant attention or intervention. Additionally, we observed frequent confusions involving function words in older students. Rather than being a cause for concern, these may reflect increased reading automaticity and speed, as short, frequent words are often overlooked by the eye during rapid reading (Rayner & Duffy, 1986; Rayner & McConkie, 1976).

In our study, various types of reading errors, as well as reading speed and accuracy, showed significant improvements from 3rd to 4th grade, with increased reading speed and accuracy alongside a reduction in errors. This pattern suggests that 4th grade serves as a key transitional period in reading development. While 1st grade is essential for establishing foundational reading skills, 4th grade appears to mark a shift from developing reading fluency to focusing on comprehension. This shift aligns with the concept of the "4th Grade Slump" described by Chall and Jacobs (2003), which identifies 4th grade as a critical stage where students transition from "learning to read" to "reading to learn" (Chall et al., 1990). Neuroimaging studies also support continued developmental changes in brain regions associated with reading automaticity, particularly in the left inferior frontal gyrus and occipitotemporal regions – areas linked to skilled reading (Shaywitz et al., 1998; Shaywitz et al., 2007). It is likely that the increased classroom demands for comprehension and analysis in 4th grade contribute to the observed improvement in basic reading tasks, as tested by the Mariette.

Study 2, dyslexia screening in a clinical setting

Method

Participants

Neuropsychologists at the Cerene testing centre (https://cerene-education.fr) included the Mariette in their assessment battery for students referred for evaluation, with families providing consent to share the data for research purposes. The sample comprised 18 children with putative neurodevelopmental deficits: 2 in 2nd grade, 3 in 3rd grade, 7 in 4th grade, and 6 in 5th grade.

Procedure

Participants were tested during the 2022–2023 school year as part of a broader neuropsychological evaluation. Parents received written information about the research and provided consent for their child's participation. Neuropsychologists followed the same standardized protocol for administering the Mariette as used in our lab.

Establishing norms. Standardized norms for reading speed and accuracy were established with the data collected in Study 1. Since reading speed and accuracy did not follow a normal distribution, norms were defined using percentiles. Since students were stopped after five minutes of reading, not all children read the same number of words during the test. To ensure meaningful error analysis, we established grade-specific cut-offs indicating the minimum number of words a child needed to read for each error type to be reliably assessed. These cut-offs, shown in Table 3, were based on the data collected in Study 1, and corresponded to the minimal number of words read by children in the 5th to 10th percentiles for either reading speed or accuracy within each grade. Errors for each error type were then calculated using only the words read up to the established cut-off. For example, in 3rd grade, a child needed to read at least 230 words for their error analysis to be relevant, and only errors made on the first 230 words were included. Given that children at the start of 1st grade typically read very few words, analysis of error types was not constructive for this group. As a result, cut-offs were established from the middle of 1st grade, as shown in Table 3, along with the number of items of each error type within these cut-offs.

We excluded from the analysis 25 students who did not reach the cut-off (8 1st graders, 4 2nd graders, 1 3rd grader, 6 4th graders, and 6 5th graders). Additionally, 62 students were excluded because their number of errors in at least one error type was more than three standard deviations away from the mean. This criterion was applied to ensure that children presenting with potential reading difficulties were not inadvertently included in the normative sample. We applied Crawford and Garthwaite's significance t-test to set the threshold number of errors for each error type beyond which a child should be considered selectively impaired. (Crawford & Garthwaite, 2007). Norms are shown in Table 4.

Data analysis. We used the same procedures as in Study 1 to assess each of the 18 children tested at the Cerene centre, calculating their reading speed, accuracy, and the number of errors for each error type. Individual results were then compared to standardized norms for

Table 3. Cut-offs for error analysis by grade.

Grade		Cut-off (number of words to be read so that errors can be analyzed)	Nb of word pairs	Nb of migratable words	Nb of pseudowords	Nb of morphologically complex words	Nb of irregular words
1st	Start	/	/	/	/	/	/
	Middle	53	0	5	3	5	4
	End	89	2	7	5	8	8
2nd	Start	106	2	8	6	10	8
	Middle	138	4	10	7	14	9
	End	153	4	11	8	17	11
3rd		230	7	14	12	23	17
4th		294	9	20	16	29	22
5th		294	9	20	16	29	22

speed, accuracy, and specific error types. For reading speed and accuracy, scores falling below the 5th percentile were considered indicative of an impairment in that specific reading process. A potential reading difficulty was identified when the number of errors for a given error type exceeded the threshold set using Crawford and Garthwaite's significant t-test (Crawford & Garthwaite, 2007).

Results

Students considered dyslexic by standard tests

The clinicians we collaborated with used various tests to screen for dyslexia, all of which assessed both reading accuracy and speed. They identified nine children in the sample as dyslexic, and the Mariette confirmed the diagnosis for eight of them. The one child (CI) not flagged as dyslexic by the Mariette had a reading speed and accuracy rate between the 10th and 15th percentile, classifying him as a slow reader but not reaching the threshold for any error type. Among the eight children with concordant diagnoses, six could not read enough words to allow for a specific error analysis. In other words, although their reading speed and accuracy indicated dyslexia according to the Mariette criteria, their reading difficulties were too severe to meet the grade-level cut-off for error analysis. For the two remaining children who did reach the cut-off for error analysis, their number of errors for each error type are shown in Table 5. Of the two, CD showed error rates exceeding the thresholds for all error types, including both consonant and vowel errors, with the exception of substitutions within function words and voicing errors. In contrast, FB displayed a selective vowel deficit (reading *mare* (*"pond"*) as *mère* (*"mother"*) or the pseudoword *klapin* as *klapion*), with a number of vowel errors 3.47 standard deviations above the control group mean, while consonant errors remained within the average range and no other error types were out of range.

Students considered non dyslexic by standard tests

Of the nine students not considered dyslexic by the clinicians, the Mariette provided a concordant conclusion for three: AFR was a slow but accurate reader and LP and IM were within the normal range for all parameters measured. However, the Mariette identified six of these students as having some type of reading difficulty. Two of these students, MC and DC, met the threshold for a mixed-dyslexia profile, with errors exceeding the thresholds in both visual processing (attentional and letter position errors for DC and attentional and other visual errors for MC) and in substitutions within function words (replacing *les* by *le*, *un* by *une*, *de* by *du*). CM exhibited a different type of mixed profile, making a lot of substitutions within function words, placing him 3.6 SD above the control mean, as well as a lot of misapplication of contextual rules (reading *cou**s**in* (*"cousin"*) as *coussin* (*"pillow"*)), placing him 3.3SD above the control mean. Finally, CE also exceed the controls for two types of errors, showing difficulties with the pronunciation of digraphs (reading *gal**et*** (*"pebble"*) as *galette* (*"galette"*)), placing him 3.6 SD above the control mean, and regularizing letter sound in irregular words (pronouncing the silent final letter "c" in *tabac* (*"tobacco"*)), placing him 2.9 SD above the control mean. The remaining two students exhibited selective deficits. LMG showed a single difficulty with reading irregular words. He made 7 regularization, placing him 3.9 SD above the mean for his peers. More precise tests focusing on the reading of irregular words would be needed to set the diagnosis of surface dyslexia. Finally, NS, made 9 substitutions within function words, placing him 4.2 SD above the control mean. Again, further testing would be needed to refine the diagnosis of a potential phonological buffer dyslexia.

Discussion

The findings from Study 2 underscore the Mariette screening tool's potential as a refined diagnostic aid

Table 4. Norms of the Mariette on speed, accuracy and different error types.

	1st graders			2nd graders					
	start	middle	end	start	middle	end	3rd graders	4th graders	5th graders
	Reading speed (words per minute)								
15th percentile	2.20	8.37	14.63	22.80	28.85	31.2	45.88	72.16	64.87
10th percentile	1.86	7.58	13.94	20.64	25.50	27.56	44.80	68.50	59.18
5th percentile	1.63	7.00	12.28	17.44	22.15	24.92	40.72	62.81	50.27
	Accuracy (%)								
15th percentile	30.06	56.92	68.11	73.35	81.15	80.06	85.46	91.45	90.82
10th percentile	24.53	54.41	65.88	67.78	78.68	77.09	84.15	89.68	89.59
5th percentile	22.16	51.10	62.13	61.78	76.65	72.45	80.71	87.70	84.32
	Error type **Threshold**, mean (SD)								
Cut-off	/	53	89	106	138	153	230	194	194
Attentional	/	**11** 3.8 (2.9)	**11** 3.3 (2.8)	**7** 2.5 (1.9)	**9** 2.9 (2.4)	**11** 3.0 (3.0)	**11** 3.3 (2.9)	**9** 2.5 (2.5)	**10** 2.7 (2.7)
Letter position	/	**5** 1.8 (1.4)	**5** 1.5 (1.4)	**5** 2.0 (1.4)	**7** 2.9 (1.9)	**8** 3.1 (2.0)	**11** 4.2 (2.5)	**10** 3.5 (2.5)	**13** 4.5 (3.2)
Other visual errors	/	**8** 2.4 (2.2)	**9** 2.6 (2.4)	**10** 2.4 (2.7)	**9** 2.3 (2.5)	**7** 2.1 (1.9)	**8** 2.7 (2.2)	**5** 1.6 (1.5)	**6** 1.9 (1.8)
Voicing	/	**3** 0.7 (0.9)	**3** 0.9 (1.0)	**3** 0.8 (0.8)	**2** 0.5 (0.7)	**2** 0.5 (0.7)	**3** 0.6 (0.9)	**2** 0.3 (0.5)	**2** 0.1 (0.3)
Vowel	/	**12** 4.2 (3.0)	**13** 4.1 (3.3)	**10** 3.6 (2.4)	**14** 4.6 (3.6)	**15** 4.6 (3.7)	**16** 5.3 (3.9)	**11** 3.7 (2.8)	**13** 4.5 (3.3)
Mispronunciation of digraphs	/	**5** 1.5 (1.4)	**3** 1.1 (1.1)	**3** 0.7 (0.8)	**3** 0.8 (1.0)	**3** 0.7 (1.0)	**2** 0.5 (0.7)	**2** 0.4 (0.7)	**2** 0.4 (0.7)
Misapplication of contextual rules	/	**5** 1.5 (1.5)	**5** 1.5 (1.6)	**4** 1.6 (1.3)	**3** 1.1 (1.0)	**3** 1.2 (1.0)	**4** 1.8 (1.2)	**4** 1.5 (1.0)	**4** 1.4 (1.1)
Regularization of irregular words	/	**5** 2.0 (1.4)	**8** 3.2 (2.1)	**5** 2.0 (1.5)	**5** 1.6 (1.3)	**5** 2.1 (1.2)	**7** 2.8 (1.8)	**5** 1.5 (1.4)	**7** 1.8 (2.0)
Function word substitutions	/	**3** 0.7 (0.8)	**4** 1.1 (1.2)	**5** 1.7 (1.5)	**7** 2.2 (1.8)	**7** 2.4 (1.8)	**11** 3.4 (2.8)	**6** 1.9 (1.7)	**7** 2.1 (1.9)
Morphological errors	/	**2** 0.1 (0.3)	**2** 0.3 (0.4)	**2** 0.4 (0.5)	**3** 0.7 (0.8)	**3** 0.8 (1.0)	**5** 1.5 (1.6)	**3** 0.9 (1.0)	**3** 0.8 (1.0)

compared to standard clinical assessments of dyslexia. For students clinically identified as dyslexic, the Mariette confirmed the diagnosis in eight out of nine cases, underscoring its reliability in alignment with standard methods. However, one exception (CI), a 3rd grader, is especially notable. This student scored between the 10th and 15th percentiles in reading speed and accuracy – a borderline profile. While clinical assessments identified this child as dyslexic, the Mariette classified them as a slow reader rather than dyslexic, suggesting it may distinguish between slow reading and true dyslexia more precisely, potentially reducing overdiagnosis. This distinction resonates with findings from other research (Catts et al., 2003; Shankweiler et al., 1999), which suggest that reading difficulties can range from isolated speed issues to full dyslexia. Further studies comparing responses on the Mariette and other diagnostic tools could help verify this hypothesis, especially in borderline cases. Another possibility could be that CI presented subtle language-based deficits not identified by the Mariette. Had we more time, follow-up tasks targeting morphosyntactic skills (e.g., oral production of clitic pronouns or homophonous definite determiners) might have clarified whether his profile aligned more with SLI or phonological buffer dyslexia, as he just reached the threshold for morphological errors.

Among the eight students with concordant diagnosis, six were unable to read enough words to allow detailed error analysis. This limitation highlights a challenge common in severe dyslexia cases, where students struggle to reach the cut-off for analyzing specific error types. For the two students who did meet the cut-off, however, the Mariette's error analysis provided meaningful differentiation: one student (CD) exhibited broad deficits across multiple dyslexia types, while the other (FB) displayed a selective vowel deficit. These results suggest that the Mariette's potential is not only in detecting dyslexia but also in identifying specific reading deficits, that could then be used to support tailored intervention strategies. This aligns with previous studies that emphasize the importance of fine-grained assessments for specific deficits to guide targeted interventions (Friedmann et al., 2010; Friedmann & Rahamim, 2014).

The results for students identified as dyslexic by the Mariette but not by clinical evaluation offer additional insights. We identified six out of nine of these students as having at least one reading difficulty, suggesting that the Mariette may be sensitive to more subtle or selective error profiles that traditional screenings might overlook. Within this group, four students met the criteria for a mixed-dyslexia profile, while the remaining

Table 5. Reading speed, accuracy and number of errors for each error type made by the patients. Described here are only the patients who passed the cut-off in terms of number of words read so that their errors could be analyzed with a higher precision. The stars indicate pathological scores relative to the standards.

			Error types									
Name, grade, diagnosis	Reading speed (words/min)	Reading accuracy (%)	Att.	Letter position	Visual	Voicing	Vowel	Digraph error	Cntxt. rule error	Reg. of irregular words	Sub. within function words	Mor. errors
CI, 3, dyslexic	45.1	85.4	6	4	5	0	7	1	2	6	8	5
CD, 4, dyslexic	46.6	79.3*	11*	16*	7*	2	26	6*	5*	11*	5	7*
FB, 2, dyslexic	19.6	70.5	6	4	8	0	12*	1	3	3	3	1
LP, 2, non dyslexic	34.6	88.7	4	4	1	1	7	1	1	4	0	3
IM, 5, non dyslexic	71.4	91.8	1	0	2	0	3	1	2	7	6	1
CM, 5, non dyslexic	95.7	85.7	8	6	4	1	9	0	5*	0	9*	2
DC, 5, non dyslexic	75.1	88.1	13*	14*	2	1	13	0	1	3	9*	2
NS, 4, non dyslexics	86.6	90.8	7	6	1	1	8	0	2	4	9*	1
LMG, 4, non dyslexic	47.3*	90.8	6	6	1	1	9	2	2	7*	3	0
AFR, 4, non dyslexic	56.1*	91.8	2	5	1	0	5	0	2	2	5	2
MC, 4, non dyslexic	59.8*	87.8	10*	10	7*	1	12	0	1	5	5	4*
CE, 3, non dyslexic	39.6*	85.3	4	6	7	1	6	3*	3	8*	1	3

two showed selective impairments. One of these students exhibited a pattern consistent with surface dyslexia, marked by a high number of regularizations of irregular words. These findings indicate that the Mariette is particularly effective at detecting surface dyslexia, even in students who perform relatively well on broader reading measures such as speed and accuracy. This sensitivity supports the importance of using tools capable of identifying specific dyslexia subtypes independently of general reading proficiency (Castles, 1996; Coltheart et al., 1983; Güven & Friedmann, 2022). Additionally, the Mariette identified one student who made significantly more substitutions of function words than controls. This pattern may suggest a primary deficit in the phonological output buffer – a hypothesis that should be tested further through oral language tasks (Dotan & Friedmann, 2015). It is important to note that these children's reading errors are not solely related to confusion with higher frequency words: on average, 44% of the responses of the 18 children tested in this study were pseudowords.

In sum, the results suggest that the Mariette offers a general assessment of reading difficulties based on reading speed and accuracy that is comparable to other standardized screeners, while providing a more detailed picture of reading difficulties through precise error analysis. This allows for the differentiation between dyslexia subtypes, which is often missed by conventional assessments. Future research should further validate the Mariette's sensitivity to specific dyslexia profiles and explore whether targeted interventions based on subtype classification can lead to improved outcomes for individuals with dyslexia.

General discussion

Our study underscores the complexity of reading development and highlights dyslexia as a heterogeneous learning disorder with multiple potential subtypes, reinforcing the need for a nuanced screening tool capable of distinguishing among these profiles across stages of reading acquisition. Study 1 demonstrated that the Mariette accurately reflects developmental trends in reading speed and accuracy, which improves steadily as children advance through primary school (Biemiller, 1977; Fuchs et al., 2001). Its strong correlation with widely used screeners reinforces its reliability (Lefavrais, 2005), while its detailed error analysis provides additional insights into specific error types – an area where it has been argued that traditional tools often fall short (Castles & Coltheart, 1993; Friedmann & Coltheart, 2018). By tracking the rates of different types of errors, the Mariette allows for the establishment of grade-based benchmarks, enhancing our understanding of typical and atypical reading patterns and supporting a more nuanced approach to dyslexia diagnosis.

In Study 2, we observed the Mariette's added sensitivity in a clinical setting, where it captured subtle error profiles missed by standard assessments. Of the 18 students assessed, only half were identified as dyslexic by the standardized tests used by clinicians, yet several of the remaining students appear to have specific deficits that went undetected. This finding echoes previous research showing that narrow definitions of dyslexia can lead to false negatives, overlooking students with more selective or atypical profiles (Kohnen et al., 2012; Potier Watkins et al., 2023). Specifically, among all 18 students tested, those whose accuracy scores fell below the 5th percentile made numerous errors across multiple error types, or they did not read enough words to conduct a detailed error analysis. Half of the students whose accuracy scores were between the 5th and 15th percentiles displayed a specific deficit, whereas all students with accuracy above the 15th percentile showed no deficit, confirming their prior non-dyslexic diagnosis.

Future research should further validate this distinction by comparing Mariette results with the tests used by clinicians to clarify where specific reading profiles might overlap or diverge. In this project, we know that clinicians did not consistently use the same tests, but we lack information on which specific tests they employed for their diagnoses.

The Mariette's success in identifying dyslexia subtypes suggests an exciting avenue for future research: examining whether tailored interventions targeting specific subtypes leads to more effective outcomes (Friedmann & Rahamim, 2014). For instance, children with phonological buffer dyslexia may benefit from interventions focused on strengthening phonological working memory, whereas children with attentional dyslexia might benefit from visual processing support, such as guided tracking or spatial attention exercises (Friedmann & Rahamim, 2014). As the multiple dyslexias theory gains traction, the Mariette provides a framework for personalized interventions, which could ultimately lead to more accurate and beneficial treatment options.

Limitations

The Mariette is only a screener. Obviously, a 5-minute test does not suffice to establish a full diagnosis of neurodevelopmental deficit, especially with a condition as complex as dyslexia. Results of the Mariette should be combined with further tests. For example, if a child makes more letter position errors than average but few attentional errors, this may suggest letter position dyslexia, which can be confirmed with additional targeted error screeners tests like the Malabi (Potier Watkins et al., 2023). Visual errors – such as letter position errors, attentional errors, or letter substitutions not attributable to attentional confusion – may arise from phonological rather than orthographic processing deficits. However, these explanations could potentially be ruled out if the children demonstrated intact word and pseudoword repetition abilities. The sharp decline in such errors over the course of first grade suggests that they likely stem from deficits in the orthographic visual analyzer rather than from impairments in the phonological output buffer. Nonetheless, oral repetition of words and pseudowords would have been necessary to confirm this interpretation. Confirmation through oral language tasks should also be considered when interpreting our analysis of the broader category of morphological errors – which, in our scoring, includes both inflectional and derivational errors. While such errors may stem from a phonological output buffer deficit, additional testing is necessary to confirm this hypothesis. Nonetheless, we argue that the error analysis performed using the Mariette provides a valuable starting point: a relatively quick yet informative tool for identifying reading profiles that merit more in-depth investigation.

Furthermore, while the Mariette provides some level of granularity about dyslexia subtypes, it does not identify all dyslexia subtypes mentioned in the introduction. For instance, neglexia, deficit in the orthographic input buffer, letter identity dyslexia or deep dyslexia, are not directly assessed by the Mariette. To accurately test for neglexia, assessments must include words where letter additions, omissions, or substitutions at word start or end create a different word (e.g., "*loin*" (far) read as "*loi*" (law)). While the Mariette does not specifically test for this, it is possible to analyze whether there is a pattern in the number of visual errors (letter position, attentional and other visual errors) at the beginning, middle, or end of words. If a pattern is found, follow-up tests can confirm whether these errors are due to neglexia, which impacts letters, or if a broader visual perception issue is at play, such as difficulties with reading numbers.

To assess the presence of a deficit in the orthographic input buffer, one can examine whether visual errors increase with word length, since impairments in the orthographic input buffer typically lead to a higher error rate for longer words. Sensitivity to word length is therefore an important diagnostic indicator. However, because the Mariette contains far more short words than long words, it may not allow for the detection of a reliable length-sensitive error pattern. A follow-up assessment using a list of words carefully controlled for length would therefore be necessary to evaluate the presence of such a deficit.

Diagnosing letter identity dyslexia typically requires specific tests that involve identifying letters in different fonts or cases. As mentioned in the introduction, it is important to distinguish this type of dyslexia from orthographic-visual analyzer dyslexia, as both disorders can cause similar reading errors. In cases where a selective visual orthographic analyzer dyslexia is detected, follow-up testing for letter identity dyslexia is recommended. This will help determine whether the child's difficulties stem from challenges in recognizing letter identity, which can present similarly but requires different interventions.

Finally, we did not include enough stimuli in this study to formally identify deep dyslexia. Deep dyslexia, though rare, requires specific tests to diagnose, which are beyond the scope of the Mariette.

Another methodological concern raised by our peers relates to the fact that we allowed children to use their finger to follow the text while reading. This decision was made based on the young age of some participants and the observation that this strategy was commonly encouraged by their teachers to help maintain their place in the text. In other words, we accepted this caveat in order to create a more natural reading context and to ensure that children had the best chance to perform at their highest level. However, this approach may have reduced the occurrence of letter position errors, particularly among the youngest children, who were the most likely to use their finger while reading (Friedmann & Rahamim, 2014). It would be valuable in future studies to confirm our findings under stricter conditions, such as by instructing all children not to use their finger to guide their reading.

In summary, the Mariette should be seen as one component of a broader set of tests used to form a comprehensive dyslexia diagnosis. Its standards help guide the interpretation of a child's reading errors, determining whether they fall within the expected range or if a higher number of errors in certain categories signals an underlying disorder, warranting further testing.

Conclusion

Our measurements show that the Mariette is a valid and reliable test for assessing reading speed, accuracy and specific errors in primary school children. What makes it so useful is the details that it provides about the nature of the errors a child makes. The developmental changes that we observed in the proportions of errors made by children according to their grade level provide practitioners and teachers with benchmarks that can help them place each pupil on a developmental scale and provide them with appropriate support.

Acknowledgements

We extend our gratitude to Clémence Ameslon, Audrey Marais, Emma Martin and Margaux Thuau for their invaluable assistance in data collection. We are deeply grateful to Elena Bombled, Amélie Guiot, Nathalie Khan and Camille Touati who conducted the testing for the children in Experiment 2 and generously shared the results of their neuropsychological assessments. We also extend our heartfelt thanks to the participating schools, whose involvement was essential to the data collection process and the success of this research. Special thanks to Professor Naama Friedmann for her insightful guidance and expertise.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by INSERM, CEA, Collège de France, Université Paris Saclay, Fondation Collège de France, and Fondation Clermont-Tonnerre. This work also benefited from the partnership with CERENE schools and their diagnostic center. The test developed as part of this research will soon be accessible via the Excello resource platform (https://excellolab.org), which will include the text, normative data for speed, success rate and error types, as well as scoring support tools.

Data availability statement

Materials and analysis code for this study are available by emailing the corresponding author.

ORCID

Marie Lubineau http://orcid.org/0000-0002-3084-5958
Cassandra Potier Watkins http://orcid.org/0000-0002-6588-0614

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

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

Al Otaiba, S., Petscher, Y., Pappamihiel, N. E., Williams, R. S., Dyrlund, 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

Awaida, M., & Beech, J. R. (1995). Children's lexical and sublexical development while learning to read. *The Journal of*

Experimental Education, *63*(2), 97–113. https://doi.org/10.1080/00220973.1995.9943815
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
Benfatto, M. N., Seimyr, GÖ, Ygge, J., Pansell, T., Rydberg, A., & Jacobson, C. (2016). Screening for dyslexia using eye tracking during reading. *PLoS One*, *11*(12), e0165508. https://doi.org/10.1371/journal.pone.0165508
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. (2004). 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
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
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
Brunsdon, R. K., Hannan, T. J., Nickels, L., & Coltheart, M. (2002). Successful treatment of sublexical reading deficits in a child with dyslexia of the mixed type. *Neuropsychological Rehabilitation*, *12*(3), 199–229. https://doi.org/10.1080/09602010244000048
Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain and Language*, *81*(1-3), 568–586. https://doi.org/10.1006/brln.2001.2548
Campbell, R., & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject : A developmental case with associated deficits of phonemic processing and awareness. *The Quarterly Journal of Experimental Psychology Section A*, *37*(3), 435–475. https://doi.org/10.1080/14640748508400944
Castles, A. (1996). Cognitive correlates of developmental surface dyslexia : A single case study. *Cognitive Neuropsychology*, *13*(1), 25–50. https://doi.org/10.1080/026432996382051
Castles, A. (2006). The dual route model and the developmental dyslexias. *London Review of Education*, *4*(1). 49–61. 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
Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, *36*(2), 151–164. https://doi.org/10.1177/002221940303600208
Chall, J. S., & Jacobs, V. A. (2003). *Poor children's fourth-grade slump*. *American Educator*.
Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis : Why poor children fall behind*. Harvard University Press.
Cohen Patrick Verstichel Stanislas Dehaene, L. (1997). Neologistic jargon sparing numbers: A category-specific phonological impairment. *Cognitive Neuropsychology*, *14*(7), 1029–1061. https://doi.org/10.1080/026432997381349
Coltheart, M. (2005). Modeling reading : The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading : A handbook* (pp. 6–23). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch1
Coltheart, M., & Kohnen, S. (2012). Acquired and developmental disorders of reading and spelling. In M. Faust (Ed.), *The handbook of the neuropsychology of language* (pp. 892–920). Wiley-Blackwell. https://doi.org/10.1002/9781118432501.ch43
Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *The Quarterly Journal of Experimental Psychology Section A*, *35*(3), 469–495. https://doi.org/10.1080/14640748308402483
Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). *DRC: A dual route cascaded model of visual word recognition and reading aloud*. 53.
Crawford, J. R., & Garthwaite, P. H. (2007). Comparison of a single case to a control or normative sample in neuropsychology : Development of a Bayesian approach. *Cognitive Neuropsychology*, *24*(4), 343–372. https://doi.org/10.1080/02643290701290146
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
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., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., 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
Diagnostic and Statistical Manual of Mental Disorders | Psychiatry Online. (s. d.). DSM Library. Consulté 2 octobre 2024, à l'adresse https://www.psychiatryonline.org/doi/book/10.1176appi.books.9780890425787
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'Ecole 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
Ellis, A. W. (1979). Developmental and acquired dyslexia : Some observations on Jorm (1979). *Cognition*, *7*(4), 413–420. https://doi.org/10.1016/0010-0277(79)90025-8
Ellis, A. W., Flude, B. M., & Young, A. W. (1987). "Neglect dyslexia" and the early visual processing of letters in words and nonwords. *Cognitive Neuropsychology*, *4*(4), 439–464. https://doi.org/10.1080/02643298708252047
Ellis, A. W., & Young, A. W. (1996). *Human cognitive neuropsychology, a textbook with readings*. Psychology Press.

Fawcett, A., & Nicolson, R. (2004). *DST-J - Dyslexia Screening Test —Junior*. Pearson.

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

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 (Eds.), *Handbook of communication disorders* (pp. 721–752). De Gruyter. https://doi.org/10.1515/9781614514909-036

Friedmann, N., & Gvion, A. (2001). Letter position dyslexia. *Cognitive Neuropsychology*, *18*(8), 673–696. https://doi.org/10.1080/02643290143000051

Friedmann, N., & Gvion, A. (2003). *TILTAN: Battery for the diagnosis of dyslexias in Hebrew*. Tel Aviv University.

Friedmann, N., & Haddad-Hanna, M. (2012). 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. (2014). Types of developmental dyslexia in arabic. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic literacy* (pp. 119–151). Springer Netherlands. https://doi.org/10.1007/978-94-017-8545-7_6

Friedmann, N., Kerbel, N., & Shvimer, L. (2010). 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., & Lukov, L. (2011). Subtypes of developmental surface dyslexia and its manifestations in Hebrew. In O. Korat & D. Aram (Eds.), *Literacy and language: Interactions, bilingualism, and difficulties* (pp. 414–444). Magnes.

Friedmann, N., & Nachman-Katz, I. (2004). Developmental neglect dyslexia in a Hebrew-reading child. *Cortex*, *40*(2), 301–313. https://doi.org/10.1016/S0010-9452(08)70125-4

Friedmann, N., & Rahamim, E. (2007). Developmental letter position dyslexia. *Journal of Neuropsychology*, *1*(2), 201–236. https://doi.org/10.1348/174866407X204227

Friedmann, N., & Rahamim, E. (2014). 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. (2001). 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

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001

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

Guggenheim, R., & Friedmann, N. (2014, juillet). *Previously undescribed type of developmental phonological dyslexia.* In International Neuropsychological Society Meeting.

Güven, S., & Friedmann, N. (2019). 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

Güven, S., & Friedmann, N. (2022). Even in predictable orthographies : Surface dyslexia in Turkish. *Scientific Studies of Reading*, *26*(6), 489–513. https://doi.org/10.1080/10888438.2022.2058399

Hindmarsh, G. P., Black, A. A., White, S. L., Hopkins, S., & Wood, J. M. (2021). Eye movement patterns and reading ability in children. *Ophthalmic and Physiological Optics*, *41*(5), 1134–1143. https://doi.org/10.1111/opo.12854

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). 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

Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S., & Zorman, M. (2010). *BALE (Batterie Analytique du Langage Ecrit)*.

Kaufman, A., & Kaufman, N. (2014). *Kaufman test of educational achievement | Third Edition*. Pearson.

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

Khentov-Kraus, L., & Friedmann, N. (2018). Vowel letter dyslexia. *Cognitive Neuropsychology*, *35*(5–6), 223–270. https://doi.org/10.1080/02643294.2018.1457517

Kohnen, S., Nickels, L., Castles, A., Friedmann, N., & McArthur, G. (2012). 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

Launay, L., Maeder, C., Roustit, J., & Touzin, M. (2018). *Evaleo 6-15*. Ortho Édition.

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. https://doi.org/10.1111/j.1467-9280.2009.02405.x

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX : A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, *36*(1), 156–166. https://doi.org/10.3758/BF03195560

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), 1404–1417. https://doi.org/10.2307/1130460

Lubineau, M., Watkins, C. P., Glasel, H., & Dehaene, S. (2024). Examining the impact of reading fluency on lexical decision results in French 6th graders. *Open Mind*, *8*, 535–557. https://doi.org/10.1162/opmi_a_00140

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

Mcarthur, G., Castles, A., Kohnen, S., Colenbrander, D., Marinus, E., Beidermann, B., Larsen, L., & Banales, E. (2023). *Motif Battery*.

McGee, R., Prior, M., Williams, S., Smart, D., & Sanson, A. (2002). The long-term significance of teacher-rated hyperactivity and reading ability in childhood: Findings from two longitudinal studies. *Journal of Child Psychology and Psychiatry*, *43*(8), 1004–1017. https://doi.org/10.1111/1469-7610.00228

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

Mousty, P., Leybaert, J., Alegria, J., Content, A., & Morais, J. (1994). BELEC. Batterie d'évaluation du langage écrit et de ces troubles. In J. Grégoire et B. Piérart (Eds.), *Evaluer les troubles de la lecture. Les nouveaux modèles théoriques et leurs implications diagnostiques* (pp. 127–145). De Boeck.

Nachman-Katz, I., & Friedmann, N. (2010). An empirical evaluation of treatment directions for developmental neglect dyslexia. *Procedia - Social and Behavioral Sciences*, *6*, 248–249. https://doi.org/10.1016/j.sbspro.2010.08.124

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

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

Potier Watkins, C., Dehaene, S., & Friedmann, N. (2023). Characterizing different types of developmental dyslexias in French: The Malabi screener. *Cognitive Neuropsychology*, *40*(7–8), 319–350. https://doi.org/10.1080/02643294.2024.2327665

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). 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

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

Rayner, K., Murphy, L. A., Henderson, J. M., & Pollatsek, A. (1989). Selective attentional dyslexia. *Cognitive Neuropsychology*, *6*(4), 357–378. https://doi.org/10.1080/02643298908253288

Scarborough, H. S. (1998). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia*, *48*(1), 115–136. https://doi.org/10.1007/s11881-998-0006-5

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

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

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., Fowler, A., Dreyer, L. G., Marchione, K. E., Shaywitz, S. E., & Shaywitz, B. A. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading*, *3*(1), 69–94. https://doi.org/10.1207/s1532799xssr0301_4

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

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., Shankweiler, D. P., Liberman, A. M., Skudlarski, P., Fletcher, J. M., Katz, L., Marchione, K. E., Lacadie, C., Gatenby, C., & Gore, J. C. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences*, *95*(5), 2636–2641. https://doi.org/10.1073/pnas.95.5.2636

Shaywitz, B. A., Skudlarski, P., Holahan, J. M., Marchione, K. E., Constable, R. T., Fulbright, R. K., Zelterman, D., Lacadie, C., & Shaywitz, S. E. (2007). Age-related changes in reading systems of dyslexic children. *Annals of Neurology*, *61*(4), 363–370. https://doi.org/10.1002/ana.21093

Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, *142*(5), 498–545. https://doi.org/10.1037/bul0000037

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*, *54*(2), 87–104. https://hal.archives-ouvertes.fr/hal-00733548

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

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'Education Nationale, de la Jeunesse et des Sports*.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*(2), 182–198. https://doi.org/10.1016/0093-934X(80)90139-X

Temple, C. M. (2006). Developmental and acquired dyslexias. *Cortex*, *42*(6), 898–910. https://doi.org/10.1016/S0010-9452(08)70434-9

Thibault, M.-P., Helloin, M.-C., & Croteau, B. (2003). Exalang – 5/8. Une batterie d'examen du langage oral et écrit chez l'enfant de 5 à 8 ans. *Travaux neuchâtelois de linguistique*, *38-39*, 129–152. https://doi.org/10.26034/tranel.2003.2587

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, 40(1), 7–26.

Traficante, D., Luzzatti, C., & Friedmann, N. (2021, août). *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., Faísca, 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., 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

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

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

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). 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., Judica, A., Orlandi, M., & Spinelli, D. (1999). Markers of developmental surface dyslexia in a language (Italian) with highgrapheme–phoneme correspondence. *Applied Psycholinguistics*, *20*(2), 191–216. https://doi.org/10.1017/S0142716499002027

Zoubrinetzky, 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