DEVELOPMENTAL DYSLEXIA: FROM CROSS-LINGUISTIC AND BILINGUAL PERSPECTIVES

EDITED BY: Fan Cao, Aaron J. Newman and Xi Becky Chen

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

DEVELOPMENTAL DYSLEXIA: FROM CROSS-LINGUISTIC AND BILINGUAL PERSPECTIVES

Topic Editors:
Fan Cao, Sun Yat-sen University, China
Aaron J. Newman, Dalhousie University, Canada
Xi Becky Chen, University of Toronto, Canada

Table of Contents

Check for
updates

Implicit Learning, Bilingualism, and Dyslexia: Insights From a Study Assessing AGL With a Modified Simon Task

Maria Vender[1]*, Diego Gabriel Krivochen[2], Beth Phillips[2], Douglas Saddy[2] and Denis Delfitto[1]

[1] Department of Cultures and Civilizations, University of Verona, Verona, Italy, [2] Centre for Integrative Neuroscience and Neurodynamics, University of Reading, Reading, United Kingdom

This paper presents an experimental study investigating artificial grammar learning in monolingual and bilingual children, with and without dyslexia, using an original methodology. We administered a serial reaction time task, in the form of a modified Simon task, in which the sequence of the stimuli was manipulated according to the rules of a simple Lindenmayer grammar (more specifically, a Fibonacci grammar). By ensuring that the subjects focused on the correct response execution at the motor stage in presence of congruent or incongruent visual stimuli, we could meet the two fundamental criteria for implicit learning: the absence of an intention to learn and the lack of awareness at the level of resulting knowledge. The participants of our studies were four groups of 10-year-old children: 30 Italian monolingual typically developing children, 30 bilingual typically developing children with Italian L2, 24 Italian monolingual dyslexic children, and 24 bilingual dyslexic children with Italian L2. Participants were administered the modified Simon task developed according to the rules of the Fibonacci grammar and tested with respect to the implicit learning of three regularities: (i) a red is followed by a blue, (ii) a sequence of two blues is followed by a red, and (iii) a blue can be followed either by a red or by a blue. Results clearly support the hypothesis that learning took place, since participants of all groups became increasingly sensitive to the structure of the input, implicitly learning the sequence of the trials and thus appropriately predicting the occurrence of the relevant items, as manifested by faster reaction times in predictable trials. Moreover, group differences were found, with bilinguals being overall faster than monolinguals and dyslexics less accurate than controls. Finally, an advantage of bilingualism in dyslexia was found, with bilingual dyslexics performing consistently better than monolingual dyslexics and, in some conditions, at the level of the two control groups. These results are taken to suggest that bilingualism should be supported also among linguistically impaired individuals.

Keywords: artificial grammar learning, implicit learning, bilingualism, dyslexia, bilingualism and dyslexia interaction

INTRODUCTION

The extent to which bilingualism can enhance executive functions (EFs) as well as metalinguistic skills (Bialystok et al., 2008, 2014) attracts vast research interest. However, a sparse number of studies has explored the interaction between bilingualism and atypical development, in order to investigate whether these advantages extend also to individuals suffering from specific impairments such as developmental dyslexia[1]. This would have a crucial social impact, since parents and teachers of impaired children often fear that bilingualism could negatively affect their linguistic development and could thus decide that one of the languages should be abandoned (Vender et al., 2018a; Garaffa et al., 2019).

Importantly, the limited available evidence seems to suggest that the positive effects associated to bilingualism in metalinguistic tasks not only extend also to bilingual children with dyslexia, but can be even more marked than in typical populations [see Vender et al. (2018b) for a study on nonword pluralization]. Conversely, the relationship between bilingualism and dyslexia in the domains of EF and implicit learning has not been examined yet.

With the aim of bridging this gap, we investigated the interaction between these two populations (bilingual and dyslexic children) in a task assessing implicit learning, using a modified Simon task in which the sequence of the stimuli is determined by the rules of an artificial grammar.

This paper is organized as follows: we first introduce the concept of artificial grammar learning (AGL), reporting the studies assessing implicit learning in bilinguals as well as in dyslexic children especially focusing on the grammar that we employed in the present study, the Fibonacci grammar. We then discuss the literature addressing the performance of bilinguals and that of dyslexics in the Simon task and formulate our research questions and predictions. Finally, we present our experimental task discussing its results and implications.

Bilingualism and Dyslexia: What Artificial Grammar Learning Can Tell Us

Artificial grammar learning is an experimental paradigm employed to investigate how sequences of symbols generated by a system are learnt. Once exposed to an artificial grammar (a set of rules that applies to an alphabet of symbols to generate strings), participants are assumed to develop some "implicit" knowledge of the regularities associated with it. In a typical AGL task, subjects first complete a training session in which they are exposed to stimuli arranged according to an invented grammar and are asked to pay attention to them, often by means of a recall task. After this training phase, they are made aware that these stimuli comply with a set of rules and are then instructed

---

[1]Developmental dyslexia is a genetic disorder characterized by a difficulty in properly acquiring reading and spelling skills, despite adequate classroom exposure, in absence of cognitive, physical, or sensorimotor impairments and socio-economical or emotional problems (Vellutino, 1979). Beyond literacy problems, dyslexia is characterized by marked linguistic deficits, affecting in particular phonological, morphological, and grammatical competence, as well as by WM and processing deficits (Vender, 2017).

to provide grammaticality judgments for new sets of items which either are consistent with these rules (i.e., grammatical) or violate them (i.e., ungrammatical).

Results of classical AGL studies (e.g., Reber, 1967), which have been extensively replicated, indicate that people are successful in discriminating grammatical from ungrammatical stimuli, although they do not display conscious knowledge of the rules. These typically remain, at least in part, implicit [see Pothos (2007) for a general review of the different theoretical accounts of AGL performance]. The ability to detect patterns and statistical regularities in an artificial grammar has been found also in very young children (Gómez and Gerken, 2000). This capacity provides evidence for statistical learning based on transitional probabilities to compute distributional information and formulate relevant hypotheses about following stimuli (Saffran et al., 1996; Gerken et al., 2005). Moreover, it correlates with natural language learning and processing (Christiansen et al., 2012), indicating that AGL can provide a useful tool for investigating the ways in which humans perceive and process stimuli, as well as for understanding higher-order cognitive functions, including language (Pothos, 2007; De Vries et al., 2008). Therefore, AGL offers new ways to investigate specific aspects of language learning that are not easily testable with natural languages, such as analyzing language acquisition and processing, while also investigating the underpinnings of the human language faculty in a controlled setting (Ettlinger et al., 2016). Using language-independent rules (which nonetheless share properties with the kind of computational devices that are hypothesized to underlie grammatical competence) and non-linguistic stimuli has several practical advantages in implicit learning paradigms: in particular, it allows speakers of different native languages to be compared across one medium (Culbertson et al., 2013); it allows young children who may not have fully acquired language as well as nonverbal populations to be tested on that medium (Gomez and Gerken, 1999); and it allows researchers to fine-tune the paradigm with a precision that is limited only by their understanding of the mathematical properties of the rules and the structures thereby generated.

There are other notable methodological benefits: the participant has not been exposed to the stimulus beforehand, so observed experimental effects can be reliably linked back to the grammar, and implicit learning can be observed independently of factors which play a major role in the natural language parsing, such as semantics and pragmatics (Lobina, 2011). More particularly, it is possible to isolate specific local units for analysis without worrying about confounding factors related to the content of the symbols being used.

Artificial grammar learning has, more recently, been used to explore implicit learning in atypical populations, including individuals suffering from language-related impairments, such as aphasia (Christiansen et al., 2010) and developmental language disorder/specific language impairment (Evans et al., 2009). As for developmental dyslexia, deficits in AGL have been reported by Pavlidou and Williams (2014), who found that school-aged children with dyslexia showed difficulties in implicit learning; more specifically, in higher-order rule-like learning. Using a nonverbal task assessing AGL by presenting geometric shapes

arranged either sequentially or in an embedded way, Pothos and Kirk (2004) found evidence for a different learning strategy in dyslexic adults in comparison to controls; impaired subjects were less skilled in processing the individual elements of the stimuli. Other studies confirmed that dyslexics are impaired in implicit learning tasks, indicating that they struggle in identifying and assimilating systematic patterns of stimuli in a structured setting, independently of the learning materials (Folia et al., 2008; Goldberg, 2014).

However, other studies have reported that dyslexics show no disadvantages in AGL (Rüsseler et al., 2006), which suggests that the complexity of the learning environment (in terms of processing costs) could play a major role (Vicari et al., 2005; Roodenrys and Dunn, 2007; Pavlidou et al., 2010; Nigro et al., 2015). Consistently, Katan et al. (2017) administered to the same group of children two AGL tasks differing in the type of grammar adopted, and found that children with dyslexia, although performing worse than controls with the grammars that, according to the authors, were more difficult to learn, showed intact learning of the less complex grammar, suggesting that they managed to extract relevant regularities from the input under less demanding conditions.

All in all, these results seem to suggest that dyslexics, despite exhibiting problems in the implicit detection and abstraction of rules under complex conditions, nevertheless do show a sensitivity to structural regularities in AGL (Pavlidou et al., 2010). Their difficulties could then be attributed to working memory (WM) restrictions: due to their limitations in WM and in processing capacity [see Nicolson and Fawcett (2008) and Vender (2017) for accounts based on processing deficits in dyslexia], dyslexics could be less efficient than their peers in formulating and simultaneously comparing different hypotheses depending on the structural regularities of the input (Baddeley, 1983).

Artificial grammar learning in bilingualism has not been extensively studied and the limited results available are mixed: Onnis et al. (2018) reported heightened performance in bilinguals in two AGL tasks while individual variables were controlled for; similarly, a bilingual advantage in statistical learning has been reported by other studies (Bartolotti et al., 2011; Escudero et al., 2016). Conversely, no differences were found by Yim and Rudoy (2013). Poepsel and Weiss (2016) compared monolingual and bilingual adults in a statistical word-learning task, reporting similar performance of the two groups with a moderate level of processing difficulty, but evidence for a bilingual advantage, with an increased level of processing difficulty, suggesting that basic statistical learning is not affected by bilingualism, whereas a bilingual advantage could arise in more complex tasks that require inhibiting potential sources of interference.

To summarize so far, the studies conducted until now have typically investigated AGL by explicitly exposing subjects to visually or auditorily presented sequences of symbols produced by a grammar, and explicitly asking subjects, after training, to provide acceptability judgments on these (or new) sequences of symbols. The results of these studies confirm that AGL takes place across different ages, measured by above-chance performance in the grammaticality tasks, in healthy subjects as well as in bilinguals, who in some cases have been found to

outperform monolinguals. Although displaying intact learning in easier conditions, dyslexic subjects have instead been found impaired in conditions requiring more costly processing.

The present study investigates AGL in monolingual and bilingual children, with and without dyslexia, using a radically different methodology: instead of overtly training the subjects with sequences of symbols and asking for grammaticality judgments after training, we administered a serial reaction time (SRT) task; more specifically, a modified version of the Simon task. In our version, the sequence of visually presented stimuli is not random, but predictable on the basis of systematic regularities that characterize the output of the grammar we used. In this way, we can fully exploit the advantages of a SRT task in order to preserve the implicit nature of AGL. Under these experimental conditions, the two main requirements for implicit learning (i.e., absence of an intention to learn and lack of awareness of the acquired knowledge) are clearly guaranteed. This constitutes an original aspect of our protocol. Even more original is our use of a set of rules belonging to a class of grammars different from those used in traditional AGL experiments, as will be discussed below.

The Fibonacci Grammar: A Simple Lindenmayer System

To date, AGL tasks have primarily used grammars in "canonical form" (Jäger and Rogers, 2012). These grammars, by definition, consist of (1) an alphabet which includes a start symbol (i.e., the symbol from which the rewriting procedure originates), rewriteable symbols (i.e., symbols which are written as other symbols, continuing the rewriting procedure), and non-rewriteable symbols (i.e., symbols that stop rewriting and correspond to the terminal forms of the strings generated) and (2) a set of rules of the form "rewrite A as B" which determine specifically how the grammar is developed by rewriting symbols in the alphabet in a stepwise manner, as will be described below. By applying these rewriting rules left-to-right sequentially to a set of symbols, grammatical "strings" are generated, also termed "words" or "sentences." An example is the kind of phrase-structure rules familiar from linguistics, where → is simply "rewrite left-hand side as right-hand side" [i.e., "every time you find the symbol in the left in your input string, replace it with the symbol(s) in the right"], follows in (1):

(1) Sentence → Noun Phrase + Verb Phrase

The rule above encodes hierarchical constituency in a sentence: a symbol Sentence is rewritten as two non-terminals Noun Phrase (NP) and Verb Phrase (VP). Further structural details can be provided in the form of the rule in (2):

(2) Noun Phrase → Determiner + Noun

In (2), both "Determiner" and "Noun" are *terminal symbols*, insofar as they do not rewrite as any other symbol. It is worth emphasizing that the second rule can only apply if the first has applied already: otherwise there is no "NP" symbol to rewrite. This strict sequentiality and inherent order in rule application is usually referred to as a "traffic convention," and it is a crucial property of phrase structure grammars.

In this respect, it should be emphasized that familiar systems of the kind that are customarily referred to in order to describe natural language structure, traditionally giving rise to the much-discussed Chomsky hierarchy Chomsky (1956), do not exhaust the landscape of rule-based formalisms.

Our implementation of AGL exploits one of these alternative formalisms: Lindenmayer systems. Lindenmayer grammars (Lindenmayer, 1968; Rozenberg and Salomaa, 1980; Prusinkiewicz and Lindenmayer, 2010) are simple deterministic recursive rewrite systems with some special properties. First, there is no distinction between nodes (nonterminals, i.e., symbols that are rewritten as other symbols; S, NP, and VP in the example above) and leaves (terminals, i.e., symbols that terminate the rewriting procedure; Determiner and Noun, above). Second, there is no "traffic convention," indicating that all expandable symbols are effectively expanded *all at once*; expansion takes place in a top-down fashion, rather than left-to-right. Finally, they present self-similarity: each generation of the grammar maps to earlier generations, such that any natural constituent of the grammar can be used to reconstruct structural context, as displayed in **Figure 1**.

An important property of L-systems is that the strings that they generate contain a systematic range of statistical regularities. These follow from the formal properties of the grammar and can be controlled and probed for without *ad hoc* modifications. As a result, stimuli generated using L-systems provide an extraordinary platform for investigating the potential and limits of statistical learning (Saddy, 2009).

As argued above, previous research has shown that humans are able to extract information from signals, including natural and artificial grammars (Shirley, 2014; Geambasu et al., 2016; Phillips, 2017). However, identifying the specific kind of operation involved in this process is controversial. A non-randomly generated signal will present surface statistical regularities locally governing the transition between distinct symbols in the string, for whichever mode of presentation under consideration. It has been shown that these surface statistical effects can be found in children as young as 8 months old (Saffran et al., 1996) as well as in other species (e.g., Fehér et al., 2017). Given a signal, a fundamental question is whether statistical mechanisms are enough for an organism to infer or learn the underlying system of rules that has generated that signal and therefore make reliable hypotheses about adjacent and non-adjacent symbols in a sequence in locally ambiguous conditions. In this context, rule learning (which requires higher-order computational operations than the calculation of immediate transition probabilities in a string) has also been shown to be available very early on and to be essential for an adequate account of language and language-like phenomena (Marcus et al., 1999; Marcus and Berent, 2003).

For the purposes of the present paper, we have used a specific L-system, a so-called Fibonacci grammar (Fib grammar henceforth)[2], defined by the following rules:

(3) $0 \rightarrow 1$
      $1 \rightarrow 1\ 0$

The interpretation of such a formalism is very simple: every instance of [0] in a sequence must be replaced by (or "rewritten as") [1], and every instance of [1] in the same sequence must be replaced by [1 0] in a top–bottom derivation. Applying these rules over and over again generates longer and longer sequences of symbols: specifically, the grammar in (3) generates derivations like the hierarchical sequence reported in **Figure 1**, where each row (a "generation" of the grammar) is a sequence of [1]s and [0]s and corresponds to a string of symbols. These strings of [1]s and [0]s can then be mapped onto linguistic or non-linguistic stimuli, across varying modalities.

An important derivational property of Fib-grammars [see Krivochen and Saddy (2018), Krivochen et al. (2018), Saddy (2018) for discussion about Fib grammars] is that each generation can be predicted if (i) we have access to the previous generation and to the rules, or (ii) we have access to two successive generations.[3]

FIGURE 1 | Self-similarity in the Fibonacci derivation.

---

[2]The Fibonacci grammar owes its name to the number of total items generated per row as well as of 0s and 1s individually, which follows the Fibonacci sequence (1, 1, 3, 5, 8, 13, . . .); see **Figure 1**.

[3]This is not surprising if we consider that the Fibonacci sequence itself is defined as a recurrence relation, where for any term F_n we have that:

(i) $F_n = F_{n-1} + F_{n-2}$

Analogously, any generation G_n can be defined by the following formula:

(ii) $G_n = G_{n-1} \,^\wedge\, G_{n-2}$

where $^\wedge$ indicates concatenation. Note that, because the relation $^\wedge$ is not commutative (i.e., does not produce an identical output regardless of the order of items, unlike the operation +), left-concatenating generation G_{n-1} to G_{n-2} does not yield the same result as right-concatenating G_{n-1} to G_{n-2} (Krivochen and Phillips, 2018). This is a non-trivial property which is essential to be aware of in order to make predictions about the symbols that come up in the string at any juncture.
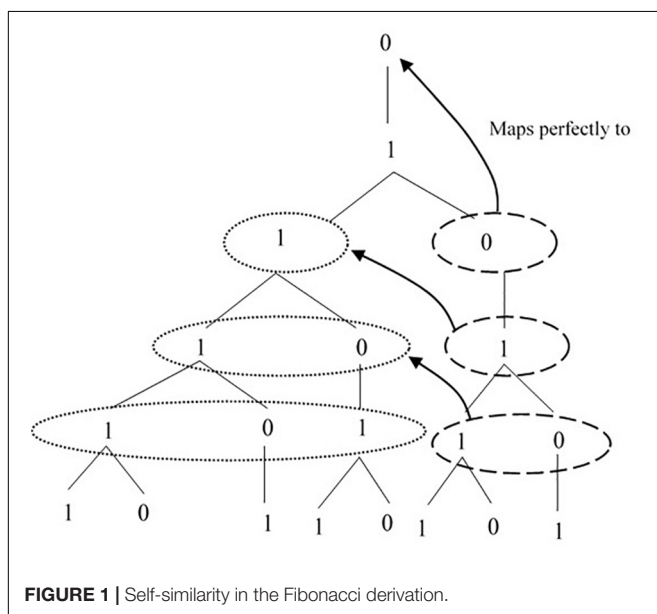
The grammar presented in (3) generates strings in which the following first-order transitional regularities hold:

(4) (a) A [0] is always followed by a [1]
     (b) Two [1] are always followed by a [0]
     (c) A single [1] can be followed by either a [0] or a [1]

These regularities imply that the following *n*-grams are never to be found in the derivation of the Fib grammar, and are thus "ungrammatical":

(4′) *00
      *111

In principle, (4c) could be seen as suggesting an element of non-determinism in the derivation of the grammar; but this is not so. The ambiguity that arises in single [1]s pertains *only* to left-to-right, very local transition probabilities: once we have more information about the string (i.e., if we have access to more symbols), these points can be disambiguated in a systematic way by reconstructing the underlying hierarchical structure (the "derivation"). In other words, simply by looking at its environment, we know without the need to reconstruct anything that if a [1] is preceded by another [1] the following symbol is [0]. If the [1] is preceded by a [0], instead, we face two possible scenarios, only one of which leads to a real ambiguity. The sequence [. . .10101], indeed, is only apparently ambiguous, since it can be disambiguated by local structure reconstruction, i.e., going back one generation: since the previous generation of [1010] is [11], and since we know that [*111] is ungrammatical, we are forced to conclude that only a [1] can complete the sequence [10101]. The only case of real ambiguity presented by the Fibonacci grammar is found in the sequence [1101], since looking back to the string alone does not provide enough information to predict what comes, as it can be followed either by a [0] or by a [1]. Here, we will not go into further details regarding structural ambiguities in the Fib-string (see Krivochen et al., 2018), but it is important to be aware of these dependencies in order to understand the type of information that is being implicitly learned.

Given the properties illustrated above, a reasonable learning hypothesis is that there are two distinct processes going on at the same time: a low-level statistical process ("low level" because it is string-based), rooted in linear relations [see regularities (4a,c) above], and a high-level process rooted in the induction of relations between non-adjacent symbols (which require us to go beyond strictly linear relations, up to phrase-structure power).

The Simon Task: Implications for Bilingualism and Dyslexia

In traditional versions of the Simon task (Simon, 1969), subjects are presented with random sequences of blue and red shapes appearing on the left or on the right side of a computer screen, and they are instructed to press distinct keys on the keyboard, depending on the color of the item only, ignoring its position on the screen. In "congruent" trials, the stimulus is on the

same side as the key to be pressed, whereas in "incongruent" trials, the correct key is on the opposite side. Performance in terms of reaction times (RTs) and accuracy is typically worse (i.e., slower RTs and lower accuracy) for incongruent trials, which require more attentional resources in order to inhibit responses based on irrelevant information (i.e., the position of the square on the screen).

It has been found that bilinguals, across different ages, are more skilled than monolinguals in tasks tapping their EFs [see Adesope et al. (2010) for a review on 63 studies investigating EF in bilinguals; but see also Hilchey and Klein (2011) and Paap (2018) for a more critical perspective on the bilingual advantage], including the Simon task: bilinguals are indeed typically faster than monolinguals in this task, on both congruent and incongruent trials (Bialystok et al., 2004, 2005; Bialystok, 2006; Morton and Harper, 2007; Martin-Rhee and Bialystok, 2008).

As for an explanation for this advantage, no consensus has been reached yet. Some scholars have proposed that bilinguals display higher inhibitory control than monolinguals (Carlson and Meltzoff, 2008; Luk et al., 2010) or better EF in general (Bialystok et al., 2004): specifically, since their two (or more) languages are always active in the brain, they need to constantly inhibit the one which is not used at a given moment. This is suggested to make them generally more adept at focusing on relevant stimuli, inhibiting irrelevant ones. However, more recent studies have suggested that attentional control, instead of inhibition and interference suppression functions, is more enhanced in bilinguals. More particularly, Zhou and Krott (2018) hypothesized that bilinguals have greater abilities in engaging and maintaining vigilant attention in task performance: this allows them to avoid temporary lapses of attention which would lead to "temporary loss of task goals from the working memory" (p. 2). Crucially, enhanced attentional control leads to better performances in both conflict and non-conflict conditions, which would explain why bilinguals' better performance in EF tasks, such as the Simon task, has been found not only in incongruent conditions, but also in congruent ones.

Conversely, EF is typically compromised in dyslexics, who display deficits in the maintenance of relevant information in WM, in both long-term memory access and retrieval and in the inhibition of irrelevant information [Varvara et al., 2014; see Booth et al. (2010) for a recent meta-analysis on children with reading disabilities].

In the present study, we tested learning of an artificial grammar by means of a modified Simon task. The paradigm was modified in two ways: (i) the sequence of stimuli was determined by the Fibonacci grammar (see section "The Fibonacci Grammar: A Simple Lindenmayer System") instead of being "randomized," and (ii) incongruent trials occurred at regular intervals (every sixth item). The first modification allowed us to verify whether statistical learning succeeds, manifesting itself in terms of faster RTs for predictable trials (corresponding to the unambiguous points in the series of visual stimuli as discussed above). The second modification, though less strictly tied to the experimental logic of the design, was implemented in order to limit the conflict between congruent and incongruent trials, by making

incongruent trials regularly occurring and therefore statistically predictable. This conflict, which plays a central role in the traditional Simon task is, for the most part, devoid of interest for the purposes of the present experimental design. Introducing a regular repeat was expected to be sufficiently easy to maintain the nature of the Simon task while adding a simple regularity for subjects to detect. Furthermore, the occurrence of the incongruent trial every 6 was long enough to allow anticipation, so as to involve some limited kind of effort. This arguably contributed to keep the task engaging for the participants.

It should be emphasized that there are important advantages in adopting the Simon task, as a widely applied experimental tool in cognitive sciences. First of all, it allows direct targeting of subjects' abilities to extract regularities from the input without conscious awareness. It also allows for the creation and presentation of stimuli which are visual instead of verbal, thus yielding a language-independent task. Furthermore, in a SRT task such as this, participants are only required to respond to visual stimuli (a challenge made relatively complex by the conflict between congruent and incongruent trials), to the effect that the participants' conscious attention is arguably diverted from the patterns that these stimuli follow. More particularly, since the participants' only concern is to respond correctly to the trials, the possibility that they take "chance" decisions is plausibly lower than in designs where they are requested to provide a grammaticality judgment, even when they feel unsure about the answer. This means that SRT paradigms are not required to meet the "zero correlation criterion" in order to observe truly implicit learning [see Dienes (2008) for a discussion about the verification of implicit learning in AGL experiments]. Evidence for implicit learning using a SRT task is provided by Cleeremans and McClelland (1991), indicating that this can offer a viable tool for assessing automatic learning of sequential material (see also Goldberg, 2014).

Research Questions and Predictions

In light of what discussed above, we were first of all interested to establish whether there was any learning of the transitional rules of the Fibonacci grammar during the execution of our modified Simon task, supported by a decrease in RTs in the trials where the following stimulus was predictable on the basis of the transitional regularities induced by the grammar on the output. Since the Fib grammar is non-canonical, arguably instantiating some kind of more abstract and potentially language-independent grammatical knowledge, this result is of interest in itself.

Second, and more importantly, we were interested to establish whether, and to what extent, these learning effects also manifested themselves within the two populations in question (bilinguals and dyslexics), and whether there was, with respect to learning, any interaction between bilingualism and dyslexia.

As for dyslexia, the sparse studies on AGL involving a SRT task (Goldberg, 2014) suggest that dyslexics may show learning improvements comparable to typically developing controls, although differences could arise in conditions requiring higher processing costs, as discussed above. Moreover, if dyslexics are found to be delayed in their learning process in comparison to controls, this could support prior predictions that the kind of

procedural knowledge involved in implicit learning is, at least to some extent, impaired in dyslexia.

As for bilingualism, although the previous results from AGL research, as seen above, are not homogeneous, we are inclined to believe that the reportedly enhanced ability of bilinguals to track distributional regularities of the input across associated representations in different languages might result in increased efficiency and flexibility in generally detecting regularities through analysis of the input (Weiss et al., 2015). Since the ability to track distributional properties in the input is most plausibly linked to unconscious procedural knowledge, it should be possible to address it with a task assessing implicit learning. Moreover, we emphasize that in the lively debate about the cognitive aspects of bi- and multilingualism [see Bialystok et al. (2008) and Costa et al. (2008) for studies reporting advantages of bilingualism; but see also Hilchey et al. (2015) and Paap et al. (2015) for more critical perspectives], the role of learning as such has received only modest attention. In the research presented here we explicitly face exactly this issue: the modified version of the Simon task that we propose here clearly provides implicit learning opportunities for the subjects.

In a nutshell, our experimental hypotheses are thus as follows: (i) we predict that, in the experimental protocol outlined here, learning should succeed for all groups, supporting the robustness of implicit learning effects in SRT paradigms and, crucially, for non-canonical grammars; (ii) we predict that differences among the three groups might also be found, with dyslexics exhibiting less efficient learning and bilinguals performing better, for the reasons just mentioned. As already emphasized, we are also particularly interested in the bilingualism/dyslexia interaction, in order to assess whether bilingualism has a positive or negative influence on the dyslexics' performance at the level of implicit learning, and whether the possible benefits of bilingualism extend also to impaired children. Based on the limited results available mentioned above, (iii) we expect in fact that the benefits of bilingualism, related to bilinguals' enhanced attentional skills and improved procedural learning skills, extend also to children with dyslexia.

THE CURRENT STUDY

Participants

Our experimental protocol was administered to 108 children divided in four groups: 30 Italian monolingual typically developing children (MC; mean age 10.0 years old, $SD = 1.2$), 30 bilingual typically developing children with Italian as an L2 (BC; mean age 10.2 years old, $SD = 1.2$), 24 Italian monolingual dyslexic children (MD; mean age 10.0 years old; $SD = 1.3$), and 24 bilingual dyslexic children (BD; mean age 10.4 years old, $SD = 1.4$).

All the monolingual children were native speakers of Italian, whereas Italian was the L2 of all the bilingual children.[4] A

---

[4]Due to the difficulties of recruiting a sufficient large sample of bilingual dyslexic children speaking the same L1, we included in our sample children with heterogeneous L1s. The L1s of the BD were: Albanian (seven children),

questionnaire was administered to gather information about their exposure to the two languages, including Age of First Exposure (AFE) to Italian, Quantity of Exposure (QE) in Italian, Traditional and Cumulative Length of Exposure (TLE and CLE) to Italian.[5] All subjects were active bilinguals using their L1 principally at home with parents and siblings and their L2 at school. The results of the questionnaire for the bilingual groups are reported in **Table 1**. No significant differences were found among the two groups concerning AFE [$t(51) = 0.504$, $p = 0.518$], QE [$t(51) = 0.612$, $p = 0.543$], TLE [$t(51) = 0.621$, $p = 0.537$], and CLE [$t(51) = 0.534$, $p = 0.595$].

Bilingual and monolingual children attended the same public schools and lived in the same areas in the north of Italy (Trento and Verona). Regarding the socio-economic status of the participants, we considered parental education, asking parents to provide information about their educational level: one point was attributed to primary education (i.e., primary and middle school), two for secondary education (i.e., high school), and three for higher education (i.e., university). Each subjects' parental education score was calculated as the average of their parents' scores. No statistically significant differences between the four groups were found [$F(3,104) = 1.558$, $p = 0.204$; see **Table 2** for mean values of each group].

Having a formal diagnosis of developmental dyslexia based on standard criteria (ICD-10, World Health Organization [WHO], 2004) was the inclusion condition for the two dyslexic groups; all the diagnostic tasks were administered in Italian, which was the language of instruction for all the children.

Finally, no children had other diagnosed or referred cognitive deficits, hearing or vision disorders, nor comorbidity with other language disorders including developmental language disorder or specific language impairment. Children were recruited through contacts with the local health system (as for part of the dyslexic children) and through the schools they were in attendance at

---

Arabic (six), Romanian (five), Spanish (two), Hindi (one), Turkish (one), Yoruba (one), and Senegalese Wolof (one). The L1s of the BC were: Romanian (nine children), Arabic (eight), Albanian (four), Hindi (two), Spanish (one), Ghanaian English (one), Yoruba (one), Moldovan (one), Serbian (one), Polish (one), and Macedonian (one).

[5]Information about the exposure to the two languages were collected by administering the Bilingual Language Exposure Questionnaire [see Unsworth et al. (2012) and Vender et al. (2016) for a description of the task and for the discussion of the concepts of TLE, a traditional index calculating subtracting the AFE of the child to their chronological age, and CLE, which is a more precise measure considering the different exposures to the two languages of the child in both the present and the past].

TABLE 1 | Means (*standard deviations*) of age of first exposure (AFE), quantity of exposure (QE), traditional length of exposure (TLE), and cumulative length of exposure (CLE) to Italian of the two bilingual groups.

	AFE (in years)	QE (in percentage)	TLE (in years)	CLE (in years)
BD	2.52	0.67	7.71	2.27
	(2.30)	(0.14)	(2.20)	(0.80)
BC	2.24	0.64	8.08	2.39
	(1.81)	(0.13)	(2.10)	(0.75)

BD, bilingual dyslexics; BC, bilingual controls.

(as for the remaining dyslexic children and all the controls); no monetary compensation was provided to participants. The study was approved by the local Ethics Committee (Department of Neurological, Biomedicine and Movement Sciences, University of Verona, Verona, Italy) and conducted in accordance with the standards specified in the 2013 Declaration of Helsinki; moreover, written informed consent was given by the parents of all the children who took part in our research study.

Materials
Preliminary Measures
All participants underwent a series of additional cognitive and linguistic tests. All children had to score within the normal ranges in the CPM Raven task measuring general intelligence (Raven et al., 1998; Italian standardization by Belacchi et al., 2008). Dyslexics had to score lower than $-2SD$ below the mean of their reference category in two out of four reading measures (measured by speed and accuracy of word and nonword reading, *Batteria per la Valutazione della Dislessia e della Disortografia Evolutiva*, by Sartori et al., 2007). Conversely, typically developing children had to score within the normal ranges. We also assessed the children's receptive vocabulary [by use of the *PPVT-R* by Dunn and Dunn (2000) Italian standardization by Stella, Pizzioli, and Tressoldi[6]], their WM (by administering the Forward and the Backward Digit Span task, FDS and BDS, of the WM test by Pickering and Gathercole, 2001) and their phonological competence [by administering a nonword repetition (NWR) task, see Vender et al. (under review)].

Modified Simon Task
The experiment was run on an Asus 15.6′ laptop using DMDX Automode version 4.3.0.1 software. The stimuli were four squares (dimensions 1012×536 pixels, BMP files) each for one of the four conditions. Each trial started with a fixation cross which appeared in the middle of the screen and remained visible for 500 ms and which was followed by a red or a blue square, either on the left or on the right side of the screen. As in traditional Simon tasks, participants were presented with four experimental conditions (blue congruent, blue incongruent, red congruent, and red incongruent) and instructed to press the number key 1 (on the left side of the keyboard) if they saw a red square and the number key 0 (on the right side of the keyboard) if they saw a blue square, irrespective of the position of the squares.

In our modification, the order of the colored squares presented to the subject was not random but instead determined by a simple deterministic recursive grammar; the Fib-grammar (described above). The strings of stimuli the grammar generates deliver a range of regularities: from simple local dependencies to higher order dependencies (as defined in section "The Fibonacci Grammar: A Simple Lindenmayer System"). From the subjects' perspective the Simon task is unchanged; however, it is possible to track the subjects' implicit learning of the regularities via RT and accuracy responses across the duration of the task.

---

[6]PPVT-R is a task addressing receptive vocabulary; children are asked to point at a picture out of an array of four alternatives to select the one that represents the word uttered by the experimenter.

TABLE 2 | Means (*standard deviations*) of the preliminary measures for each group.

	Nonverbal Intelligence Raven[a]	Vocabulary PPVT-R[b]	Word reading speed[a]	Nonword reading speed[a]	Word reading accuracy[a]	Nonword reading accuracy[a]	Forward digit span	Backward digit span	Nonword repetition	Parental education
BD	0.12	90.50	−2.13	−0.88	−2.71	−2.75	23.92	10.33	0.60	1.92
	(*0.58*)	(*0.13*)	(*2.18*)	(*1.45*)	(*1.32*)	(*1.10*)	(*3.47*)	(*3.59*)	(*0.13*)	(*0.42*)
MD	0.10	108.21	−3.75	−2.86	−2.20	−2.13	24.87	9.50	0.67	2.00
	(*0.75*)	(*13.18*)	(*2.73*)	(*2.58*)	(*1.80*)	(*1.42*)	(*3.69*)	(*4.23*)	(*0.12*)	(*0.25*)
BC	0.20	95.80	0.99	0.65	0.04	0.22	27.57	13.60	0.84	1.85
	(*0.83*)	(*13.50*)	(*0.77*)	(*0.68*)	(*0.95*)	(*0.80*)	(*6.63*)	(*5.24*)	(*0.11*)	(*0.33*)
MC	0.37	105.13	0.25	0.26	0.25	0.26	29.23	12.97	0.86	2.03
	(*0.74*)	(*9.09*)	(*0.87*)	(*0.80*)	(*0.87*)	(*0.80*)	(*4.77*)	(*3.92*)	(*0.07*)	(*0.36*)

[a]*Z-scores.* [b]*Standard scores; other scores are raw scores. BD, bilingual dyslexics; MD, monolingual dyslexics; BC, bilingual controls; MC, monolingual controls.*

Both accuracy and RTs were collected: each item remained on the screen for 1000 ms if there was no response before the next item was shown. Participants were asked to answer as quickly and accurately as possible. The timing started with the onset of the item and ended with the response of the subject. There were eight random practice trials in which subjects received feedback; after the training, they had the chance to ask questions before the experiment began. The modified Simon task comprised three blocks of 144 trials each, for a total of 432 stimuli, and took 10–15 min to complete.

As discussed in the introduction, the Fib-grammar comprises two rules, which converted into the colored stimuli are:

(5) red → blue (i.e., 0 → 1)
    blue → blue, red (i.e., 1 → 1, 0)

First of all, we wanted to verify whether there were improvements related to learning the following first-order transitional regularities:

 (i) a red is always followed by a blue (a sequence of two reds is ungrammatical);
 (ii) two blues must be followed by a red (a sequence of three blues is similarly ungrammatical), and
 (iii) a blue can be followed by a red or a blue.

Moreover, in order to be sure that these improvements were related to the learning of these regularities and not to a general effect of habituation to the task, we compared performance in ambiguous (unpredictable) and unambiguous (predictable) items.

It must be noticed that, due to the formal properties of the grammar, as reviewed above, blue items were more frequent than red ones. Finally, as in every Simon task, both congruent and incongruent items were tested: unlike in traditional Simon tasks, however, the incongruent trial occurred every sixth item, for the reasons discussed above (see section "The Simon Task: Implications for Bilingualism and Dyslexia").

To summarize, we employed this modified Simon task to identify differences in performance between monolingual and bilingual healthy and dyslexic children with the aim of assessing their ability to unconsciously pick up the regularities of the Fib-grammar.

First, we examined whether all groups successfully learned the regularities in (4a–c): the fact that a red is always followed by a blue was expected to be the easiest to acquire (section "Analysis 1: Blue Items Occurring After Red Items"). That two blues are followed by a red was instead predicted to be more difficult, since the memory load was higher: to succeed in this task, it is not sufficient to consider the item which has just appeared, but it is necessary to remember also the one occurring immediately before it (section "Analysis 2: Red Items Occurring After Two Blue Items"). Finally, to verify whether any improvements across blocks found in the previous analyses were really determined by the learning of the relevant regularities, and not by a general effect of habituation to the task, we compared RTs and accuracy in unambiguous trials (determined by 6a,b) and ambiguous ones (see 6c); lower or no improvement was expected in the ambiguous condition, where subjects could not benefit from learning the regularities delivered by the grammar (as discussed above) in predicting the color of the upcoming item (section "Analysis 3: Predictable vs. Unpredictable Items").

Procedure

All children were tested individually in a quiet room by the first author. They were administered the preliminary tasks followed by the modified Simon task. The Simon task lasted approximately 10–15 min, with a short break after the end of the second block. The whole experimental session lasted approximately 60 min (45 min for the preliminary tasks and 10–15 min for the Simon task).

RESULTS

Preliminary Measures

Mean and SDs of each group in each preliminary task are reported in **Table 2**.

Results of the preliminary measures were analyzed by carrying out a series of one-way ANOVAs with group (MC, BC, MD, and BD) as the independent variable and performance in each task as a dependent variable. The four groups did not differ in age [$F(3,104) = 0.720$, $p = 0.542$] nor in general nonverbal intelligence [$F(3,104) = 1.135$, $p = 0.339$]. Conversely they differed in reading

measures, including speed of word reading [$F(3,104) = 33.249$, $p < 0.001$], accuracy of word reading [$F(3,104) = 39.335$, $p < 0.001$], speed of nonword reading [$F(3,104) = 28.830$, $p < 0.001$], and accuracy of nonword reading [$F(3,104) = 49.773$, $p < 0.001$]. *Post hoc* comparisons with Bonferroni correction (*post hoc* comparisons henceforth) revealed that in word reading MD were slower than BC, MC ($p < 0.001$), and BD ($p < 0.05$), who were in turn slower than both BC and MC ($p < 0.001$); no differences were found between MC and BC ($p = 1.000$). Moreover, MD and BD were less accurate than BC and MC ($p < 0.001$); no differences were found between MD and BD ($p = 0.939$), neither between MC and BC ($p = 1.000$). As for nonwords, MD were slower than all other groups ($p < 0.001$), whereas BD were slower than BC ($p < 0.01$) and MC ($p < 0.05$). MC and BC performed similarly ($p = 1.000$); moreover, MD and BD were less accurate than MC and BC ($p < 0.001$); no differences were found between MD and BD ($p = 1.000$) and MC and BC ($p = 1.000$).

Differences were reported also in PPVT-R [$F(3,104) = 11.163$, $p < 0.001$]; as shown by *post hoc* comparisons, BD showed a poorer vocabulary in comparison to MD and MC ($p < 0.001$), whereas BC scored lower than MD ($p < 0.01$) and MC ($p < 0.05$). No differences were found between MD and MC ($p = 1.000$) and between BD and BC ($p = 0.706$).

Significant differences were also found for both FDS [$F(3,104) = 6.593$, $p < 0.001$] and BDS [$F(3,104) = 5.624$, $p < 0.01$]. *Post hoc* comparisons showed that in FDS, MD scored lower than MC ($p < 0.01$) but similarly to BC ($p = 0.292$), whereas BD scored lower than both MC ($p < 0.001$) and BC ($p < 0.05$). No differences were found between MD and BD ($p = 1.000$) nor between MC and BC ($p = 1.000$). As for BDS, instead, MD performed more poorly than MC ($p < 0.05$) and BC ($p < 0.01$), whereas BD had lower BDS scores than BC ($p < 0.042$) but not than MD ($p < 0.292$). No differences were found between MD and BD nor between MC and BC ($p = 1.000$).

Group differences were found also in NWR [$F(3,104) = 34.680$, $p < 0.001$]; as revealed by *post hoc* comparisons, both MD and BD performed worse than MC and BC ($p < 0.001$), whereas they performed similarly to each other ($p = 0.327$); also MC and BC performed similarly ($p = 1.000$).

Summarizing, the two dyslexic groups differed significantly from the control groups in all literacy measures, in WM tasks, and in phonological competence, whereas no differences were found in nonverbal intelligence and receptive vocabulary. The resulting profile is consistent with the typical cognitive and linguistic profile of children with dyslexia. Differences in vocabulary, but not in literacy, WM, and phonological competence are instead in line with the literature describing the typical profile of bilingual children (Bialystok et al., 2010). Since receptive vocabulary is reported to be relatively unimpaired in dyslexia (Vender et al., 2017), the fact that bilingual controls underperformed monolingual dyslexics and that no negative effects of dyslexia were observed should not be surprising.

Modified Simon Task

In order to compare the performance of the four groups of children in the modified Simon task, both RTs and accuracy rates

were considered. RTs were calculated only for correct answers, representing 93.59% of the responses. Answers given earlier than 200 ms, corresponding to 1.2% of the trials, were excluded from the analysis since they might reflect anticipatory response prior to proper stimulus processing. As outlined above, there was a time limit for participants' responses, since the items disappeared after 1000 ms if no key was pressed; non-responses corresponded to 4.3% of the trials. All remaining trials were within the interval defined by the 2.5SDs intra-subject average, and thus no data were considered outliers. We then calculated the mean RT of each participant in each of the conditions tested.

In order to provide an answer to our research questions, aiming to verify whether participants showed evidence of having learnt the regularities of the input and whether group differences emerged, three distinct analyses were performed. In section "Analysis 1: Blue Items Occurring After Red Items," the learning of the first regularity (a red is always followed by a blue) was investigated, whereas the fact that two blues are always followed by a red was assessed in section "Analysis 2: Red Items Occurring After Two Blue Items." Finally, in section "Analysis 3: Predictable vs. Unpredictable Items," we compared blue items being entirely predictable based on statistical regularities (blues occurring after a red) with those being completely unpredictable (blues occurring after a sequence of blue–blue–red–blue, which was ambiguous and could be followed by either a blue or a red, as discussed in section "The Fibonacci Grammar: A Simple Lindenmayer System"). This final analysis was particularly useful to verify whether improvements in the task were really dependent on the learning of the relevant regularities: if no differences between predictable and unpredictable trials were found, improvements could indeed be related to a general effect of habituation to the task, and not to implicit learning.

Analysis 1: Blue Items Occurring After Red Items

To verify whether children learnt that a blue item always appeared after a red one, we analyzed responses to all congruent and incongruent blue trials following a red one, comparing RTs and accuracy rates of the four groups of participants across the three blocks of stimuli. As shown in **Tables 3**, **4**, reporting mean RTs and accuracy rates for each group in each block and condition, all groups displayed a decrease in RTs from Block 1 to Block 3, both in congruent and in incongruent trials. Moreover, bilingual dyslexics are faster than the other groups in each condition, whereas monolingual dyslexics were the slowest. As for accuracy, it was at ceiling for all groups in the congruent conditions, whereas it was lower in the incongruent trials, especially for the monolingual dyslexics.

We ran a repeated-measures ANOVA with *Bilingualism* and *Dyslexia* as between-subject variables and *Congruency* (Congruent vs. Incongruent trials) and *Block* (1, 2, and 3) as within-subject variables.

As for RTs, we found a main effect of *Bilingualism* [$F(1,104) = 5.521$, $p < 0.05$, $\eta_p^2 = 0.051$], no main effect of *Dyslexia* [$F(1,104) = 0.011$, $p = 0.916$, $\eta_p^2 = 0.000$], and no *Bilingualism* × *Dyslexia* interaction [$F(1,104) = 0.729$, $p = 0.395$, $\eta_p^2 = 0.007$], indicating that bilinguals are faster than monolinguals in processing blue items occurring after

TABLE 3 | Mean (*standard deviation*) reaction times (RTs) in each condition for each group ("Analysis 1: Blue Items Occurring After Red Items").

	C_1	C_2	C_3	I_1	I_2	I_3
BD	463.98	453.66	415.90	662.77	638.84	614.80
	(73.47)	(62.76)	(70.89)	(85.30)	(80.96)	(95.61)
MD	506.92	492.17	456.06	700.45	670.73	660.53
	(71.91)	(67.06)	(62.41)	(82.47)	(110.16)	(97.25)
BC	489.94	471.39	424.96	672.23	650.49	611.65
	(65.58)	(64.90)	(60.17)	(83.88)	(105.10)	(112.70)
MC	479.59	480.50	437.86	695.48	700.25	637.17
	(75.08)	(79.61)	(70.97)	(93.29)	(95.09)	(91.45)

C, congruent; I, incongruent; 1, Block 1; 2, Block 2; 3, Block 3; BD, bilingual dyslexics; MD, monolingual dyslexics; BC, bilingual controls; MC, monolingual controls.

TABLE 4 | Mean (*standard deviation*) accuracy in each condition for each group ("Analysis 1: Blue Items Occurring After Red Items").

	C_1	C_2	C_3	I_1	I_2	I_3
BD	0.97	0.96	0.97	0.79	0.79	0.82
	(0.04)	(0.04)	(0.04)	(0.19)	(0.18)	(0.21)
MD	0.97	0.92	0.96	0.77	0.71	0.80
	(0.07)	(0.16)	(0.06)	(0.17)	(0.25)	(0.22)
BC	0.99	0.97	0.98	0.79	0.80	0.81
	(0.02)	(0.04)	(0.03)	(0.21)	(0.17)	(0.14)
MC	0.99	0.99	0.99	0.88	0.92	0.92
	(0.02)	(0.02)	(0.02)	(0.13)	(0.12)	(0.10)

C, congruent; I, incongruent; 1, Block 1; 2, Block 2; 3, Block 3; BD, bilingual dyslexics; MD, monolingual dyslexics; BC, bilingual controls; MC, monolingual controls.

a red one, irrespective of dyslexia. *Block* was also significant [$F(1,104) = 43.415$, $p < 0.001$, $\eta_p^2 = 0.297$], while the other interactions were not significant. This indicates that all groups showed an improvement in RTs across the task: specifically, RTs were faster in Block 2 than in Block 1 ($p < 0.05$) and in Block 3 than in Block 2 ($p < 0.001$). *Congruency* was also significant [$F(1,104) = 966.322$, $p < 0.001$, $\eta_p^2 = 0.904$], with congruent items being processed faster than incongruent ones. No interaction was significant, indicating that improvements were reported for both congruent and incongruent trials and for all groups.

As for accuracy, instead, we found a main effect of *Dyslexia* [$F(1,104) = 11.047$, $p < 0.01$, $\eta_p^2 = 0.096$], no main effect of *Bilingualism* [$F(1,104) = 0.883$, $p = 0.350$, $\eta_p^2 = 0.008$], and a significant *Bilingualism × Dyslexia* interaction [$F(1,104) = 8.255$, $p < 0.01$, $\eta_p^2 = 0.074$], indicating that the negative effect of dyslexia was limited to the monolingual children, with bilingual dyslexics performing more accurately than monolingual dyslexics and similarly to the two groups of controls. In this case, neither *Block* was significant [$F(1,104) = 2.020$, $p = 0.135$, $\eta_p^2 = 0.019$], nor the relevant interactions, indicating that no improvement in accuracy was found across the blocks in any of the groups.

Congruency was instead significant [$F(1,104) = 192.397$, $p < 0.001$, $\eta_p^2 = 0.649$], as well as the interaction *Congruency × Dyslexia* [$F(1,104) = 3.920$, $p = 0.050$, $\eta_p^2 = 0.036$]

and the interaction *Congruency × Bilingualism × Dyslexia*: [$F(1,104) = 7.181$, $p < 0.01$, $\eta_p^2 = 0.065$], whereas *Congruency × Bilingualism* was not significant. To understand this interaction, we ran two separate two-way ANOVAs with *Bilingualism* and *Dyslexia* as fixed factors and mean RT in congruent trials or in incongruent trials as dependent variables. When considering congruent trials we found a significant effect of *Dyslexia* [$F(1,104) = 9.449$, $p < 0.01$, $\eta_p^2 = 0.083$], no effect of *Bilingualism* [$F(1,104) = 0.181$, $p = 0.672$, $\eta_p^2 = 0.002$], and no interaction between them [$F(1,104) = 2.247$, $p = 0.137$, $\eta_p^2 = 0.021$], whereas when considering incongruent trials we found a significant effect of *Dyslexia* [$F(1,104) = 8.336$, $p < 0.01$, $\eta_p^2 = 0.074$], no effect of *Bilingualism* [$F(1,104) = 1.699$, $p = 0.195$, $\eta_p^2 = 0.016$], but a significant interaction between them [$F(1,104) = 8.627$, $p < 0.004$, $\eta_p^2 = 0.077$], indicating that in incongruent trials monolingual dyslexics were less accurate than bilingual dyslexics, who performed similarly to the two control groups.

As these results show, all groups prove to have acquired the relevant regularity, showing increasingly lower RTs across the blocks. However, group differences were found: bilinguals were overall faster than monolinguals, and monolingual dyslexics were less accurate than the other groups, especially with incongruent items. Data point thus to the presence of a positive effect of bilingualism in dyslexia: bilingual dyslexics, indeed, were overall more accurate than their monolingual peers, and less disturbed by the presence of incongruent trials. The difference between monolingual and bilingual dyslexics was more evident in more complex conditions, in which higher processing costs are arguably required.

Analysis 2: Red Items Occurring After Two Blue Items

To assess the learning of the regularity predicting that two blues are always followed by a red and the presence of group differences, we considered all red items occurring after a sequence of two blues, comparing performance of the four groups across the three blocks, while distinguishing congruent and incongruent trials. Mean RTs and accuracy rates are reported in **Tables 5**, **6**. In this case as well, all groups showed a decrease in RTs from Block 1 to Block 3; as in the previous analysis, monolingual dyslexics were the slowest, while bilinguals (both dyslexics and controls) were faster. All groups were more accurate in the congruent than in the incongruent conditions, with dyslexics being generally less accurate then controls.

As for RTs, we found a main effect of *Dyslexia* [$F(1,104) = 3.863$, $p < 0.05$, $\eta_p^2 = 0.051$], a marginally significant effect of *Bilingualism* [$F(1,104) = 5.378$, $p = 0.052$, $\eta_p^2 = 0.037$], and no interaction [$F(1,104) = 0.083$, $p = 0.773$, $\eta_p^2 = 0.001$], indicating that dyslexics were slower than controls, and that bilinguals tended to be faster than monolinguals.

Congruency was significant [$F(1,104) = 561.869$, $p < 0.001$, $\eta_p^2 = 0.848$], with incongruent items being processed more slowly than congruents. There was also a *Congruency × Dyslexia* interaction [$F(1,104) = 12.947$, $p < 0.001$, $\eta_p^2 = 0.114$], while the other interactions were not significant. Considering mean RTs in the whole task, we found that dyslexics were slower than controls

TABLE 5 | Mean (*standard deviation*) RTs (in ms) in each condition for each group ("Analysis 2: Red Items Occurring After Two Blue Items").

	C_1	C_2	C_3	I_1	I_2	I_3
BD	528.16	546.49	499.90	673.46	686.13	669.20
	(70.58)	(73.52)	(65.09)	(88.34)	(107.40)	(132.18)
MD	569.47	566.88	534.42	708.33	701.03	703.35
	(58.422)	(69.52)	(72.04)	(101.76)	(103.91)	(85.99)
BC	531.51	532.75	491.25	634.39	634.70	609.51
	(63.87)	(65.40)	(59.47)	(100.54)	(101.41)	(93.11)
MC	551.39	553.10	512.23	646.68	682.91	620.45
	(76.47)	(89.07)	(78.23)	(103.00)	(96.45)	(120.53)

C, congruent; I, incongruent; 1, Block 1; 2, Block 2; 3, Block 3; BD, bilingual dyslexics; MD, monolingual dyslexics; BC, bilingual controls; MC, monolingual controls.

TABLE 6 | Mean (*standard deviation*) accuracy in each condition for each group ("Analysis 2: Red Items Occurring After Two Blue Items").

	C_1	C_2	C_3	I_1	I_2	I_3
BD	0.93	0.89	0.90	0.77	0.64	0.60
	(0.05)	(0.07)	(0.11)	(0.16)	(0.28)	(0.27)
MD	0.93	0.85	0.86	0.83	0.65	0.72
	(0.07)	(0.17)	(0.11)	(0.17)	(0.30)	(0.25)
BC	0.96	0.93	0.92	0.82	0.80	0.69
	(0.07)	(0.07)	(0.08)	(0.21)	(0.23)	(0.29)
MC	0.97	0.94	0.95	0.79	0.81	0.76
	(0.05)	(0.07)	(0.06)	(0.23)	(0.18)	(0.22)

C, congruent; I, incongruent; 1, Block 1; 2, Block 2; 3, Block 3; BD, bilingual dyslexics; MD, monolingual dyslexics; BC, bilingual controls; MC, monolingual controls.

with incongruent trials [$t(106) = 3.156$, $p < 0.01$], but not with congruent trials [$t(106) = 0.966$, $p = 0.366$].

Block was also significant [$F(1,104) = 17.160$, $p < 0.001$, $\eta_p^2 = 0.145$]; specifically significant differences were found between Block 2 and Block 3 ($p < 0.001$), but not between Block 1 and Block 2 ($p = 0.543$). No other significant interactions were found, indicating that improvements in RTs were equally reported in all groups and for both congruent and incongruent trials.

As for accuracy, we found a main effect of *Dyslexia* [$F(1,104) = 8.249$, $p < 0.01$, $\eta_p^2 = 0.073$], no effect of *Bilingualism* [$F(1,104) = 0.619$, $p = 0.433$, $\eta_p^2 = 0.037$], and no interaction [$F(1,104) = 0.002$, $p = 0.961$, $\eta_p^2 = 0.000$], indicating that dyslexics were less accurate than controls, irrespective of bilingualism.

Congruency was significant [$F(1,104) = 159.283$, $p < 0.001$, $\eta_p^2 = 0.605$], indicating lower accuracy for incongruent trials for all groups, as testified by the absence of significant interactions. *Block* was also significant [$F(1,104) = 17.160$, $p < 0.001$, $\eta_p^2 = 0.145$], as well as the interaction *Block × Dyslexia* [$F(1,104) = 5.755$, $p = 0.004$, $\eta_p^2 = 0.052$], *Congruency × Block* [$F(1,104) = 3.779$, $p < 0.05$, $\eta_p^2 = 0.035$], and *Congruency × Block × Dyslexia* [$F(1,104) = 3.082$, $p < 0.05$, $\eta_p^2 = 0.029$]. Paired sample *t*-tests separated for *Dyslexia* (dyslexics vs. controls) revealed that with congruent trials both

groups showed a decrease in performance between Blocks 1 and 2 [dyslexics: $t(47) = 3.451$, $p < 0.01$; controls: $t(59) = 3.809$, $p < 0.001$], but not between Blocks 2 and 3 [dyslexics: $t(47) = 0.556$, $p = 0.581$; controls: $t(59) = 0.527$, $p = 0.600$]. As for incongruent trials, instead, dyslexics showed a decline between 1 and 2 [$t(47) = 3.992$, $p < 0.001$], and not between 2 and 3 [$t(47) = 0.330$, $p = 0.743$], whereas on the contrary controls showed a decline between Blocks 2 and 3 [$t(59) = 2.576$, $p < 0.01$], but not between Blocks 1 and 2 [$t(59) = 0.061$, $p = 0.951$]. This indicates that in the most complex condition (with the incongruent trials), dyslexics became inaccurate earlier than controls.

To sum up, all groups showed an improvement in RTs in correspondence to the red trials following a sequence of two blues, considering both congruent and incongruent trials. However, dyslexics were generally slower, especially with incongruent trials, whereas bilinguals tended to be faster. Concerning accuracy, instead, dyslexics made generally more errors than controls, irrespective of bilingualism, and all groups had more problems with incongruent stimuli. Moreover, accuracy decreased across the task, arguably as an effect of fatigue, especially for dyslexics who seem to be affected by tiredness earlier than controls.

Analysis 3: Predictable vs. Unpredictable Items

To verify whether the improvements in speed found across blocks in the previous analyses were really determined by the learning of the relevant regularities, and not by a general effect of habituation to the task, we compared RTs and accuracy of the four groups in predictable and unpredictable trials across the three blocks. For this purpose, we compared performance in items being unpredictable (where the blue trials followed a blue–blue–red–blue sequence and were thus uncontroversially ambiguous from the perspective of string-based statistical regularities, as discussed in section "The Fibonacci Grammar: A Simple Lindenmayer System"), and in the predictable items considered in section "Analysis 2: Red Items Occurring After Two Blue Items" (blue trials following a red). Since unpredictable items never occurred in correspondence to an incongruent trial, we considered only congruent items for this analysis. As can be noted in **Tables 7, 8**, responses to predictable items are generally faster and more accurate (ceiling performance) than those to ambiguous ones for all groups. As in the previous analysis, bilinguals are faster than monolinguals, irrespective of dyslexia, whereas both groups of dyslexics tend to be less accurate than controls.

We ran a repeated-measures ANOVA with *Bilingualism* and *Dyslexia* as between-subject variables and *Predictability* (Predictable vs. Unpredictable) and *Block* (1, 2, and 3) as within-subject variables.

As for RTs, we found a main effect of *Bilingualism* [$F(1,104) = 4.765$, $p < 0.05$, $\eta_p^2 = 0.044$], no main effect of *Dyslexia* [$F(1,104) = 0.488$, $p = 0.486$, $\eta_p^2 = 0.005$], and no *Bilingualism × Dyslexia* interaction [$F(1,104) = 1.308$, $p = 0.255$, $\eta_p^2 = 0.013$], indicating that bilinguals are generally faster than monolinguals, irrespective of dyslexia.

Predictability was also significant [$F(1,104) = 236.710$, $p < 0.001$, $\eta_p^2 = 0.697$], with predictable items yielding faster

TABLE 7 | Mean (*standard deviation*) RTs (in ms) in each condition for each group ("Analysis 3: Predictable vs. Unpredictable Items").

	P_1	P_2	P_3	U_1	U_2	U_3
BD	463.98	453.66	415.90	535.15	532.43	515.68
	(73.47)	(62.76)	(70.89)	(79.67)	(87.42)	(92.70)
MD	511.50	496.51	457.62	589.82	562.35	541.89
	(69.86)	(65.01)	(63.33)	(79.33)	(74.74)	(84.80)
BC	489.93	471.39	424.96	529.92	539.20	493.88
	(65.58)	(64.90)	(60.17)	(68.08)	(72.97)	(59.83)
MC	479.59	480.50	437.86	553.61	542.84	530.77
	(75.08)	(79.61)	(70.97)	(93.12)	(92.39)	(92.04)

P, predictable; U, unpredictable; 1, Block 1; 2, Block 2; 3, Block 3; BD, bilingual dyslexics; MD, monolingual dyslexics; BC, bilingual controls; MC, monolingual controls.

TABLE 8 | Mean (*standard deviation*) accuracy in each condition for each group ("Analysis 3: Predictable vs. Unpredictable Items").

	P_1	P_2	P_3	U_1	U_2	U_3
BD	0.97	0.96	0.97	0.92	0.94	0.89
	(0.04)	(0.04)	(0.04)	(0.10)	(0.07)	(0.13)
MD	0.96	0.92	0.96	0.92	0.83	0.90
	(0.07)	(0.16)	(0.06)	(0.10)	(0.25)	(0.13)
BC	0.99	0.97	0.98	0.96	0.94	0.94
	(0.01)	(0.04)	(0.03)	(0.09)	(0.08)	(0.10)
MC	0.99	0.99	0.99	0.97	0.96	0.95
	(0.02)	(0.02)	(0.02)	(0.04)	(0.06)	(0.08)

P, predictable; U, unpredictable; 1, Block 1; 2, Block 2; 3, Block 3; BD, bilingual dyslexics; MD, monolingual dyslexics; BC, bilingual controls; MC, monolingual controls.

RTs than unpredictable ones. This held for all groups, as testified by the absence of significant interactions. *Block* was also significant [$F(1,104) = 41.946$, $p < 0.001$, $\eta_p^2 = 0.289$], while the other interactions were not significant. This indicates that all groups showed an improvement in RTs across the task. *Predictability* × *Block* was also significant [$F(1,104) = 5.306$, $p < 0.01$, $\eta_p^2 = 0.049$], while the other interactions were not. Paired samples *t*-tests revealed a significant improvement in RTs from Block 1 to Block 2 for predictable items [$t(107) = 2.248$, $p < 0.05$] but not for unpredictables [$t(106) = 1.048$, $p = 0.297$]; both unpredictable and predictable items, instead, were processed faster in Block 3 than in Block 2 [respectively, $t(107) = 9.428$, $p < 0.001$; and $t(106) = 3.461$, $p < 0.01$].

Regarding accuracy, instead, which was overall very high for all groups and especially for predictable items, we found a main effect of *Dyslexia* [$F(1,104) = 10.801$, $p < 0.01$, $\eta_p^2 = 0.094$], no main effect of *Bilingualism* [$F(1,104) = 0.238$, $p = 0.627$, $\eta_p^2 = 0.002$], and no *Bilingualism* × *Dyslexia* interaction [$F(1,104) = 2.465$, $p = 0.120$, $\eta_p^2 = 0.023$], indicating that dyslexics were less accurate than controls.

Predictability was significant [$F(1,104) = 54.813$, $p < 0.001$, $\eta_p^2 = 0.345$], but not its interactions, indicating that predictable items were processed more accurately (with almost 100% accuracy for all groups) than unpredictable ones by all groups.

Block was also significant [$F(1,104) = 3.671$, $p < 0.05$, $\eta_p^2 = 0.034$], as well as the interaction *Block* × *Bilingualism* × *Dyslexia* [$F(1,104) = 5.123$, $p < 0.01$, $\eta_p^2 = 0.047$]. No other interaction was significant. To understand the interaction, we ran a series of paired samples *t*-tests comparing general accuracy in Blocks 1, 2, and 3 in all four groups. We found that monolingual dyslexics performed worse in Block 1 than in Block 2 [$t(23) = 2.281$, $p < 0.05$], but similarly in Blocks 2 and 3 [$t(23) = 1.484$, $p = 0.151$]. Bilingual dyslexics showed instead the opposite trend, performing similarly in Blocks 1 and 2 [$t(23) = 0.893$, $p = 0.381$], but worse in Block 3 than in Block 2 [$t(23) = 2.095$, $p < 0.05$]. The two groups of controls showed instead a similar performance in all blocks [bilingual controls, Blocks 1–2: $t(29) = 1.818$, $p = 0.079$, Blocks 2–3: $t(29) = 0.367$, $p = 0.716$; monolingual controls, Blocks 1–2: $t(29) = 0.650$, $p = 0.521$, Blocks 2–3: $t(29) = 0.771$, $p = 0.447$]. This seems to indicate that, although both groups of dyslexics become generally more inaccurate throughout the task, which could be again an effect of fatigue, monolingual dyslexics seemed to be affected by tiredness earlier than bilingual dyslexics.

Summarizing, results show that, although RTs decreased for both predictable and unpredictable items, the improvement was significantly higher for the predictable items, indicating that it must be due to the learning of the relevant rules. This was also confirmed by the fact that accuracy was higher in predictable items. Notice moreover that the absence of interactions with predictability indicates that group differences, with bilinguals being faster and dyslexics being less accurate, held for both cases.

DISCUSSION

In this study, we assessed learning of an artificial grammar in monolingual and bilingual children, with and without a diagnosis of dyslexia, by means of a modified Simon task in which the order of the stimuli was not random but determined by the Fibonacci grammar.

As emphasized in section "Research Questions and Predictions," we were interested in investigating (i) whether there was implicit learning of the regularities characterizing the Fibonacci grammar and (ii) whether group differences emerged, especially in relation to the interaction between bilingualism and dyslexia. To address these research questions, we ran three separate analysis, comparing the performance of the four groups in learning that a red is always followed by a blue (section "Analysis 1: Blue Items Occurring After Red Items") and that two blues are always followed by a blue (section "Analysis 2: Red Items Occurring After Two Blue Items"). To be sure that improvements were really related to the learning of these statistical regularities, and not to a general effect of habituation to the task, we also compared the blues following a red, which were completely predictable, to the blues following the sequence of blue–blue–red–blue, which were instead unpredictable (section "Analysis 3: Predictable vs. Unpredictable Items").

Although group differences were found, with bilinguals being always faster than monolinguals and dyslexics less accurate than controls, as will be discussed below, it is worth emphasizing

that all groups showed evidence of implicit learning, as clearly confirmed by shorter RTs and improved accuracy found in unambiguous trials, which could be correctly foreseen once these regularities were learnt. In ambiguous trials, instead, the impossibility to rely on local transition probabilities prevented participants to perform as fast and accurately as with the predictable ones. Although RTs decreased for both types of trial, as a possible effect of habituation to the task, we found that the improvements in RTs and accuracy were significantly higher for the unambiguous trials, suggesting that learning had occurred. Moreover, improvements in unambiguous trials were found as early as between Blocks 1 and 2, but only between Blocks 1 and 3 for ambiguous trials. This indicates that the learning of the regularities yielded by the Fibonacci grammar took place relatively early and, in fact, before the appearance of the habituation effect to the Simon task. Finally, group effects were similar across ambiguous and unambiguous trials, with bilinguals exhibiting faster RTs and dyslexics lower accuracy.

Given these general learning effects, we further verified whether each of the two first-order transitional regularities [see 4(a–c)] had been learnt. According to the first regularity, red trials could only be followed by blue ones: results confirmed that this regularity was successfully acquired by all groups, as showed by increasingly shorter RTs, with differences being detected as early as between Blocks 1 and 2. Importantly, this improvement was found for both congruent and incongruent trials, with responses to the latter being slower and less accurate. As for accuracy, we found a negative effect of dyslexia limited to the monolingual children: specifically, bilingual dyslexics were more accurate in reacting to incongruent trials than monolingual dyslexics, and as accurate as the two control groups. This suggests that bilingualism could confer an advantage to the impaired children in the most difficult experimental conditions.

We observed that learning also took place for the second regularity, according to which a sequence of two blues must be followed by a red: again, this was observed in both congruent and incongruent trials for all groups, who showed decreased RTs between Blocks 2 and 3, suggesting that this regularity was acquired at a later stage than the first one. This is arguably related to its higher complexity, which requires participants to consider not only the immediate predecessor of the current stimulus, but also the preceding one. In this case as well, group differences were found; dyslexics were slower, especially in incongruent trials, and also less accurate than controls, whereas bilinguals tended to be faster than monolinguals. In this case, we also found a decrease in accuracy: all groups, despite being faster in predicting the occurrence of a red trial after two blues, became less accurate as the task progressed. This is arguably an effect of fatigue, particularly evident in this more difficult condition.

Summarizing, our findings lead to the important conclusion that all groups of subjects, including the children suffering from dyslexia, were able to learn the first-order regularities characterizing the Fibonacci grammar used, generated as a specific instantiation of a Lindenmayer system and assessed by means of a modified Simon task.

The other crucial focus of our work lied in the analysis of the effects of bilingualism and dyslexia in this task: interestingly, we found that bilinguals, both dyslexics and controls, were always faster than monolinguals in reacting to the stimuli appearing on the screen, for both congruent and incongruent trials. This points to a generalized bilingual advantage, consistently with other studies reviewed in the introduction and reporting shorter response times by bilinguals in the Simon task. Importantly, our results point to an extension of the advantages of bilingualism also to impaired children, indicating that bilingualism could be beneficial for dyslexics, who in some cases even performed at the level of the monolingual controls (Analysis 2: Red Items Occurring After Two Blue Items), at least in the domain of EFs and controlled attention. Conversely, dyslexics, including both monolinguals and bilinguals, were generally less accurate than controls, indicating that they struggled more than their peers with the Simon task. This result is in line with our expectations too: as argued in the literature and discussed above dyslexia can also be characterized in terms of a processing inefficiency, leading to reduced processing and memory resources available to impaired children, as well as to lower levels of controlled attention and interference suppression. This is also compatible with the fact that poorer responses were more marked in the presence of items requiring more complex processing (incongruent trials) and thus, arguably, more effortful to learn. These results confirm our expectations about group differences in the task, with dyslexics showing difficulties arguably due to their processing or memory limitations. Bilinguals, on the contrary, displayed an advantage over monolinguals which, interestingly, extended to impaired subjects, and which could be interpreted as reflecting bilinguals' increased abilities in tasks requiring controlled attention.

To sum up, our results prompt two interesting considerations, related to the novelty of our protocol and to our research questions. First, on the one side, we extended the results that have been obtained with grammars traditionally employed in the AGL literature. Our results show that learning of an artificial grammar takes place even with a generative system that instantiates more abstract, and relatively language-independent, grammatical knowledge. On the other side, we demonstrated that learning of grammar-induced regularities can be detected with a modified Simon task, which has the advantage, of maximizing the elimination of residual explicit learning and metarepresentational awareness effects that are often found in AGL investigation. More particularly, in such SRT paradigms, the subjects are never explicitly made aware of being involved in potential grammatical learning. Firstly, they are distracted from paying attention to the statistical regularities in the succession of the visual stimuli, since they have to cope with the cognitive challenge represented by the potential asymmetry of location between visual stimulus and motor response. Secondly, as is generally the case for SRT tasks, subjects are never asked about the potential learning outcome, which could be objectively detected, in our protocol, in terms of increased reduction of RT for the predictable trials with respect to the unpredictable ones, besides the generalized RT reduction that can be interpreted as an effect of habituation to the task. Therefore, our results convincingly show that the observed learning must have taken place implicitly, while subjects were focused on an entirely different task (correctly reacting to blue and red squares irrespective of the location at which

they appear on the screen) and are therefore throughout the whole process unaware of analyzing potential regularities in the sequence of items.

Second, as for the existence of group differences our data point to a general bilingual advantage in terms of RTs and to a general dyslexic disadvantage in terms of accuracy in the task. As discussed above, the shorter RTs of bilinguals can be attributed to their enhanced attentional control and specifically to their ability to maintain high levels of attention in performing the task, whereas the difficulties exhibited by dyslexics can arguably be attributed to their lower processing resources. Crucially, the bilingual advantage has also been found in impaired children: bilingual dyslexics consistently performed better than the monolingual dyslexics, reaching the accuracy levels of the two control groups in the acquisition of the easiest regularity (predicting that a red is always followed by a blue). This result suggests that bilingualism does not produce negative effects in dyslexics, as is sometimes erroneously believed; on the contrary, it can lead to significant cognitive and linguistic advantages.

Finally, this bilingual advantage is found in the familiar domain of attentional control and inhibitory skills and cannot easily be directly attributed to enhanced performance at the level of implicit learning. As repeatedly emphasized, our results show that implicit learning took place for all groups involved, crucially including (monolingual) dyslexics. In fact, as a measure of methodological caution, it must be acknowledged that all group differences we detected concerned both ambiguous and unambiguous trials, to the effect that it is difficult to disentangle the cognitive effects induced by the Simon task from those linked to the implicit learning task. We leave this issue to future research. A natural follow up could be that of administering subjects, besides our modified Simon task, a traditional Simon task, in which the sequence of the items is really random, in order to evaluate the emergence of group differences based on direct comparison between the measurement of group effects in implicit learning and the measurement of group effects in EF enhancement.

Another exciting direction of development aims at disentangling the effects of implicit learning that may be exclusively rooted in the computation of statistically based transitional probabilities from the (possible) effects that stem from the subject's capacity to assign a hierarchical structure, given the sequences generated by the Fib-grammar. As discussed in the introduction the latter is a necessary condition that must be met in order to perform above-chance in the choice of the following symbol when presented with a sequence *blue–red–blue–red–blue* (i.e., 10101). These local configurations differ in constituency structure with respect to the local sequence *blue–blue–red–blue* (i.e., 1101), which we have used in the present study to define string-based *real* points of *ambiguity* [see Krivochen et al. (2018) for formal discussion].

In this way, the methodological advantages of our modified Simon task could be made relevant not only for measuring and evaluating learning differences among populations, but also for assessing the precise nature of implicit learning and discriminating between different accounts of implicit learning.

CONCLUSION

In this experiment, implicit learning of an artificial grammar in monolingual and bilingual children with and without dyslexia was investigated by means of a modified Simon task (a specific instance of SRT task) in which the sequence of stimuli followed the rules of a Fib-grammar (one of the Lindenmayer systems). Results clearly support the idea that learning took place, since participants of all groups became increasingly sensitive to properties of the input manifested by local sequences of red and blue items. Importantly, the two low-level regularities that we assessed [in (4a–b); i.e., a red is followed by a blue, and two blues are followed by a red] were acquired by all groups; however, overall group differences were found, with bilinguals being faster than monolinguals, and dyslexics less accurate than controls. These results, besides pointing toward some new exciting avenues of research, as discussed above, already clearly indicate that the benefits of bilingualism crucially extend to impaired children, suggesting that bilingualism should be encouraged and supported also in linguistically impaired individuals.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The study was approved by the local Ethics Committee (Comitato Etico del Dipartimento di Scienze Neurologiche e del Movimento dell'Università degli Studi di Verona; "Ethic Committee of the Department of Neurological and Movement Sciences at the University of Verona") and conducted in accordance with the standards specified in the 2013 Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

DS conceived and developed the modified Simon task. DD and MV conceived the whole experimental protocol. MV collected the data, ran the statistical analyses, and wrote the manuscript. DK and BP contributed to the writing of section "The Fibonacci Grammar: A Simple Lindenmayer System". All authors contributed to the interpretation of the results, revised the work critically for important intellectual content, and gave the final approval of the version to be published.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803

Baddeley, A. D. (1983). Working memory. *Philos. Trans. R. Soc. Lond.* 302B, 311–324.

Bartolotti, J., Marian, V., Schroeder, S., and Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Front. Cogn.* 2:324. doi: 10.3389/fpsyg.2011.00324

Belacchi, C., Scalisi, T. G., Cannoni, E., and Cornoldi, C. (2008). *Manuale CPM. Coloured Progressive Matrices. Standardizzazione Italiana*. Firenze: Giunti OS Organizzazioni Speciali.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79. doi: 10.1037/cjep2006008

Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A., et al. (2005). Effect of bilingualism on cognitive control in the simon task: evidence from MEG. *Neuroimage* 24, 40–49. doi: 10.1016/j.neuroimage.2004.09.044

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I. M., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859

Bialystok, E., Luk, G., Peets, K. F., and Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Biling. Lang. Cogn.* 13, 525–531. doi: 10.1017/s1366728909990423

Bialystok, E., Peets, K. F., and Moreno, S. (2014). Producing bilinguals through immersion education: development of metalinguistic awareness. *Appl. Psycholinguist.* 35, 177–191. doi: 10.1017/S0142716412000288

Booth, J. N., Boyle, J. M., and Kelly, S. W. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: findings from a meta-analysis. *Br. J. Dev. Psychol.* 28, 133–176. doi: 10.1348/026151009x485432

Carlson, S., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x

Chomsky, N. (1956). Three models for the description of language. *IRE Trans. Inform. Theory* 2, 113–124. doi: 10.1109/tit.1956.1056813

Christiansen, M. H., Conway, C. M., and Onnis, L. (2012). Similar neural correlates for language and sequential learning: evidence from event-related brain potentials. *Lang. Cogn. Process.* 27, 231–256. doi: 10.1080/01690965.2011.606666

Christiansen, M. H., Kelly, M. L., Shillcock, R. C., and Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition* 116, 382–393. doi: 10.1016/j.cognition.2010.05.015

Cleeremans, A., and McClelland, J. (1991). Learning the structure of event sequences. *J. Exp. Psychol. Gen.* 120, 235–253. doi: 10.1037//0096-3445.120.3.235

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013

Culbertson, J., Smolensky, P., and Wilson, C. (2013). Cognitive biases, linguistic universals, and constraint-based grammar learning. *Top. Cogn. Sci.* 5, 392–424. doi: 10.1111/tops.12027

De Vries, M. H., Monaghan, P., Knecht, S., and Zwitserlood, P. (2008). Syntactic structure and artificial grammar learning: the learnability of embedded hierarchical structures. *Cognition* 107, 763–774. doi: 10.1016/j.cognition.2007.09.002

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Prog. Brain Res.* 168, 49–64.

Dunn, L. M., and Dunn, L. M. (2000). "Peabody–test di vocabolario recettivo," in *Adattamento Italiano e Standardizzazione a Cura di*, eds G. Stella, C. Pizzioli, and P. E. Tressoldi (Torino: Omega Edizioni).

Escudero, P., Mulak, K. E., Fu, C. S., and Singh, L. (2016). More limitations to monolingualism: bilinguals outperform monolinguals in implicit word learning. *Front. Psychol.* 7:1218. doi: 10.3389/fpsyg.2016.01218

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., and Wong, P. C. M. (2016). The relationship between artificial and second language learning. *Cogn. Sci.* 40, 822–847. doi: 10.1111/cogs.12257

Evans, J. L., Saffran, J. R., and Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *J. Speech Lang. Hear. Res.* 52, 321–335. doi: 10.1044/1092-4388(2009/07-0189)

Fehér, O., Ljubièiæ, I., Suzuki, K., Okanoya, K., and Tchernichovski, O. (2017). Statistical learning in songbirds: from self-tutoring to song culture. *Philos. Trans. R. Soc. B* 372:20160053. doi: 10.1098/rstb.2016.0053

Folia, V., Uddén, J., Forkstam, C., Ingvar, M., Hagoort, P., and Peterssona, K. M. (2008). Implicit learning and dyslexia. *Ann. N. Y. Acad. Sci.* 1145, 132–150.

Garaffa, M., Vender, M., Sorace, A., and Guasti, M. T. (2019). Is it possible to differentiate multilingual children and children with developmental language Disorder? *Lang. Soc. Pol.* doi: 10.17863/CAM.37928

Geambasu, A., Ravignani, A., and Levelt, C. C. (2016). Preliminary experiments on human sensitivity to rhythmic structure in a grammar with recursive self-similarity. *Front. Neurosci.* 10:281. doi: 10.3389/fnins.2016.00281

Gerken, L., Wilson, R., and Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *J. Child Lang.* 32, 249–268. doi: 10.1017/s0305000904006786

Goldberg, N. (2014). *Imprints of Dyslexia: Implicit Learning and the Cerebellum*. Netherlands: Utrecht institute for linguistics OTS. Doctoral dissertation.

Gomez, R., and Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109–135. doi: 10.1016/s0010-0277(99)00003-7

Gómez, R., and Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends Cogn. Sci.* 4, 178–186. doi: 10.1016/s1364-6613(00)01467-4

Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on non-linguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658. doi: 10.3758/s13423-011-0116-7

Hilchey, M. D., Saint-Aubin, J., and Klein, R. M. (2015). "Does bilingual exercise enhance cognitive fitness in traditional non- linguistic executive processing tasks?," in *The Cambridge Handbook of Bilingual Processing*, ed. J. Schwieter (Cambridge: Cambridge University Press), 586–613. doi: 10.1017/cbo9781107447257.026

Jäger, G., and Rogers, J. (2012). Formal language theory: refining the chomsky hierarchy. *Philos. Trans. R. Soc. B* 367, 1956–1970. doi: 10.1098/rstb.2012.0077

Katan, P., Kahta, S., Sasson, A., and Schiff, R. (2017). Performance of children with developmental dyslexia on high and low topological entropy artificial grammar learning task. *Ann. Dyslexia* 67, 163–179. doi: 10.1007/s11881-016-0135-1

Krivochen, D., and Phillips, B. (2018). A model for a lindenmayer reconstruction algorithm. *arXiv*

Krivochen, D., Phillips, B., and Saddy, D. (2018). *Classifying Points in Lindenmayer Systems: Transition Probabilities and Structure Reconstruction*. Available at: https://www.researchgate.net/publication/329365577_Classifying_points_in_Lindenmayer_systems_transition_probabilities_and_structure_reconstruction_v_11 (accessed January 20, 2019).

Krivochen, D., and Saddy, D. (2018). Towards a classification of lindenmayer systems. *arXiv*

Lindenmayer, A. (1968). Mathematical models for cellular interactions in development I. Filaments with one-sided inputs. *J. Theor. Biol.* 18, 280–299. doi: 10.1016/0022-5193(68)90079-9

Lobina, D. (2011). Recursion and the competence/performance distinction in AGL tasks. *Lang. Cogn. Process.* 26, 1563–1586. doi: 10.1080/01690965.2011.560006

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., and Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: response inhibition versus interference suppression. *Brain Cogn.* 74, 347–357. doi: 10.1016/j.bandc.2010.09.004

Marcus, G. F., and Berent, I. (2003). Are there limits to statistical learning? *Science* 300, 53–54.

Marcus, G. F., Vijayan, S., Bandi Rao, S., and Vishton, P. M. (1999). Rule-learning in seven-month-old infants. *Science* 283, 77–80. doi: 10.1126/science.283.5398.77

Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. Lang. Cogn.* 11, 81–93. doi: 10.1017/s1366728907003227

Morton, J. B., and Harper, S. N. (2007). What did simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x

Nicolson, R. I., and Fawcett, A. J. (2008). *Dyslexia, Learning and the Brain.* Cambridge, MA: The MIT Press.

Nigro, L., Jiménez-Fernández, G., Simpson, I. C., and Defior, S. (2015). Implicit learning of non-linguistic and linguistic regularities in children with dyslexia. *Ann. Dyslexia* 66, 202–218. doi: 10.1007/s11881-015-0116-9

Onnis, L., Chun, W. E., and Lou-Magnuson, M. (2018). Improved statistical learning abilities in adult bilinguals. *Biling. Lang. Cogn.* 21, 422–433.

Paap, K. P. (2018). "The bilingual advantage debate: quantity and quality of the evidence," in *The Handbook of the Neuroscience of Multilingualism*, ed. J. W. Schwieter (Hoboken, NJ: Wiley-Blackwell).

Paap, K. P., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014

Pavlidou, E. V., Kelly, L. M., and Williams, J. M. (2010). Do children with developmental dyslexia have impairments in implicit learning? *Dyslexia* 16, 143–161. doi: 10.1002/dys.400

Pavlidou, E. V., and Williams, J. M. (2014). Implicit learning and reading: insights from typical children and children with developmental dyslexia using the artificial grammar learning (AGL) paradigm. *Res. Dev. Disabil.* 35, 1457–1472. doi: 10.1016/j.ridd.2014.03.040

Phillips, B. (2017). *Symmetry of Hierarchical Artificial Grammars and Its Effects on Implicit Learning.* Reading: University of Reading. MSc dissertation.

Pickering, S. J., and Gathercole, S. E. (2001). *Working Memory Test Battery for Children.* London: Pearson Assessment.

Poepsel, T. J., and Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition* 152, 9–19. doi: 10.1016/j.cognition.2016.03.001

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychol. Bull.* 133, 227–244. doi: 10.1037/0033-2909.133.2.227

Pothos, E. M., and Kirk, J. (2004). Investigating learning deficits associated with dyslexia. *Dyslexia* 10, 61–76. doi: 10.1002/dys.266

Prusinkiewicz, P., and Lindenmayer, A. (2010). *The Algorithmic Beauty of Plants*, 2nd Edn. New York, NY: Springer-Verlag.

Raven, J., Court, J. H., and Raven, J. C. (1998). *Raven Manual, Section 1 (General overview) and Section 2 (Coloured Progressive Matrices).* Oxford: Oxford Psychologist Press.

Reber, A. S. (1967). Implicit learning of artificial grammars. *J. Verb. Learn. Verb. Behav.* 6, 855–863. doi: 10.1016/s0022-5371(67)80149-x

Roodenrys, S., and Dunn, N. (2007). Unimpaired implicit learning in children with developmental dyslexia. *Dyslexia* 14, 1–15. doi: 10.1002/dys.340

Rozenberg, G., and Salomaa, A. (1980). *The Mathematical Theory of L-Systems.* New York: Academic Press.

Rüsseler, J., Gerth, I., and Münte, T. F. (2006). Implicit learning is intact in adult developmental dyslexic readers: evidence from the serial reaction time task and artificial grammar learning. *J. Clin. Exp. Neuropsychol.* 28, 808–827. doi: 10.1080/13803390591001007

Saddy, J. D. (2009). "Perceiving and processing recursion in formal grammars," in *Paper Presented at Recursion: Structural Complexity in Language and Cognition Conference at the University of Massachusetts (Amherst)*, (Amherst, MA: University of Massachusetts).

Saddy, J. D. (2018). *Structure, Constituency, and Labelling.* North Ryde, AU: Macquarie University.

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928. doi: 10.1126/science.274.5294.1926

Sartori, G., Job, R., and Tressoldi, P. E. (2007). *Batteria per la Valutazione Della Dislessia e della Disortografia Evolutiva-2 [Battery for the Assessment of Dyslexia and Developmental Dysorthography-2].* Firenze: Giunti OS Organizzazioni Speciali.

Shirley, E. J. (2014). *Representing and Remembering Lindenmayer-Grammars.* Ph.D. dissertation, University of Reading, Reading

Simon, J. R. (1969). Reactions towards the source of stimulation. *J. Exp. Psychol.* 81, 174–176. doi: 10.1037/h0027448

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., and Tsimpli I. (2012). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Appl. Psycholinguist.* 33, 1–41. doi: 10.1017/S0142716412000574

Varvara, P., Varuzza, C., Sorrentino, A. C. P., Vicari, S., and Menghini, D. (2014). Executive functions in developmental dyslexia. *Front. Hum. Neurosci.* 8:120. doi: 10.3389/fnhum.2014.00120

Vellutino, F. R. (1979). *Dyslexia: Theory and Research.* Cambridge, MA: MIT Press.

Vender, M. (2017). *Disentangling Dyslexia: Phonological and Processing Impairment in Developmental Dyslexia.* Frankfurt: Peter Lang.

Vender, M., Garraffa, M., Sorace, A., and Guasti, M. T. (2016). How early L2 children perform on Italian clinical markers of SLI: a study of clitic production and nonword repetition. *Clin. Linguist. Phon.* 30, 150–169. doi: 10.3109/02699206.2015.1120346

Vender, M., Hu, S., Mantione, F., Delfitto, D., and Melloni, C. (2018a). The production of clitic pronouns: a study on bilingual and monolingual dyslexic children. *Front. Psychol.* 9:2301. doi: 10.3389/fpsyg.2018.02301

Vender, M., Hu, S., Mantione, F., Savazzi, S., Delfitto, D., and Melloni, C. (2018b). Inflectional morphology: evidence for an advantage of bilingualism in dyslexia. *Int. J. Biling. Educ. Biling.* 1–8. doi: 10.1080/13670050.2018.1450355

Vender, M., Mantione, F., Savazzi, S., Delfitto, D., and Melloni, C. (2017). Inflectional morphology in dyslexia: italian children's performance in a nonword pluralization task. *Ann. Dyslexia* 67, 401–426. doi: 10.1007/s11881-017-0152-8

Vicari, S., Finzi, A., Menghini, L., Marrota, L., Baldi, S., and Petrosini, L. (2005). Do children with developmental dyslexia have an implicit learning deficit. *J. Neurol. Neurosurg. Psychiatry* 76, 1392–1397. doi: 10.1136/jnnp.2004.061093

Weiss, D. J., Poepsel, T., and Gerfen C. (2015). "Tracking multiple inputs: the challenge of bilingual statistical learning," in *Implicit and Explicit Learning of Languages*, ed. P. Rebuschat (Amsterdam: John Benjamins Press), 167–190.

World Health Organization [WHO] (2004). *International Statistical Classification of Diseases and Health Related Problems (ICD-10*, 2nd Edn. Geneva: World Health Organization.

Yim, D., and Rudoy, J. (2013). Implicit statistical learning and language skills in bilingual children. *J. Speech Lang. Hear. Res.* 56, 310–322. doi: 10.1044/1092-4388(2012/11-0243)

Zhou, B., and Krott, A. (2018). Bilingualism enhances attentional control in non-verbal conflict tasks – evidence from ex-Gaussian analyses. *Biling. Lang. Cogn.* 21, 162–180. doi: 10.1017/s1366728916000869

Check for updates

Developmental Letter Position Dyslexia in Turkish, a Morphologically Rich and Orthographically Transparent Language

Selçuk Güven[1,2] and Naama Friedmann[3]*

[1] School of Communication Sciences and Disorders, McGill University, Montreal, QC, Canada, [2] Department of Speech and Language Therapy, Anadolu University, Eskişehir, Turkey, [3] Language and Brain Lab, Sagol School of Neuroscience and School of Education, Tel Aviv University, Tel Aviv, Israel

We present the first report of a specific type of developmental dyslexia in Turkish, letter position dyslexia (LPD). LPD affects the encoding of letter positions, leading to letter migrations within words. In a multiple case study of 24 Turkish-speaking children with developmental LPD, we examined in detail the characteristics of this dyslexia and explored its manifestation in Turkish. We used migratable words, in which a migration creates another existing word (e.g., signer-singer), which exposed the migration errors of the participants. In sharp contrast with the common assumption that dyslexics in transparent languages, including Turkish, do not make reading errors, we have shown that right stimuli can detect even up to 30% reading errors. The participants made migrations in reading aloud, comprehension, lexical decision, and same-different tasks, in both words and non-words. This indicates that their deficit is in the orthographic-visual analysis stage, a stage that precedes the orthographic input lexicon and is shared by the lexical and non-lexical routes. Their repetition of non-words and migratable words was normal, indicating that their phonological output stages are intact. They did not make digit migrations in reading numbers, indicating that the orthographic-visual analyzer deficit is orthographic-specific. The properties of Turkish allowed us to examine two issues that bear on the cognitive model of reading: consonant-consonant transpositions were far more frequent than consonant-vowel and vowel-vowel migrations. This indicates that the orthographic-visual analyzer already classifies letters into consonants and vowels, before or together with letter position encoding. Furthermore, Turkish is very rich morphologically, which has allowed us to examine the effect of the morphological structure of the target word on migrations. We found that there was no morphological effect on migrations: morphologically complex words did not yield more (nor fewer) migrations than monomorphemic ones, migrations crossed morpheme boundaries and did not preserve the morphological structure of the target word. This suggests that morphological analysis follows the letter-position encoding stage.

Keywords: developmental dyslexia, letter position dyslexia, Turkish, transparent orthography, morphology

INTRODUCTION

Dyslexia is a general term relating to a deficit in reading. By now, more than 20 different types of dyslexia have been reported, each with different error types and different characteristics, each resulting from impairments in different stages of the word reading process (Marshall, 1984; Castles and Coltheart, 1993; Ellis and Young, 1996; Jackson and Coltheart, 2001; Castles, 2006; Coltheart and Kohnen, 2012; Castles et al., 2014; Friedmann and Haddad-Hanna, 2014; Hanley, 2017; Friedmann and Coltheart, 2018). Even though much work has been done about types of dyslexia in various languages, almost no studies have examined the way different types of dyslexia manifest themselves in Turkish. In fact, only one paper reported a specific type of dyslexia in Turkish, and it was a study of acquired dyslexia (Raman and Weekes, 2005); we could not find any study that described specific types of developmental dyslexia in Turkish.

In this study, we describe, for the first time, a specific type of developmental dyslexia in Turkish and its characteristics. We report on a multiple case study of 24 Turkish-speaking children who show a developmental dyslexia that affects their ability to encode the position of letters: letter position dyslexia (LPD).

According to the dual-route model (Coltheart, 1981, 1985, 1987; Patterson, 1981; Ellis et al., 1987; Shallice, 1988; Humphreys et al., 1990; Coltheart et al., 1993, 2001; Ellis and Young, 1996), the early stage of word reading includes the visual analysis of the letter string. It encodes the abstract identity of the letters and their relative positions (Coltheart, 1981; Humphreys et al., 1990; Ellis and Young, 1996). Letter position dyslexia is a deficit in the function that encodes the relative positions of letters within words, which leads to letter migrations within words (e.g., *slime → smile, cloud → could)*. Studies in several languages found that LPD is mainly manifested in migratable words, i.e., words in which letter migration creates other existing words (such as *flies* and *files*, *slept* and *spelt*). Most of the migration errors affect the middle letters, whereas the first and the last letters are relatively immune to migrations. Errors in migratable words are affected by frequency so that more errors occur when the target word (*spelt*) is less frequent than its migration counterpart (*slept*) (Friedmann and Rahamim, 2007). Individuals with LPD were found to make significantly more migrations that involve only consonants than migrations of a vowel and a consonant (Khentov-Kraus and Friedmann, 2018).

Letter position dyslexia has been reported so far for Hebrew, Arabic, and English (Friedmann and Gvion, 2001, 2005; Friedmann and Rahamim, 2007, 2014; Friedmann et al., 2010; Friedmann and Haddad-Hanna, 2012, 2014; Kohnen et al., 2012), in both acquired form (Friedmann and Gvion, 2001, 2005, for Hebrew; and Friedmann and Haddad-Hanna, 2012 for Arabic) and developmental form (Friedmann and Gvion, 2005; Friedmann and Rahamim, 2007; Friedmann et al., 2010 for Hebrew; Friedmann and Haddad-Hanna, 2012, 2014, for Arabic; Kohnen et al., 2012 for English). Until now, LPD has not been reported for Turkish.

A Brief Description of the Characteristics of the Turkish Language and Orthography

Turkish is morphologically very rich, and a single Turkish word may include multiple suffixes (e.g., the word "*güldüremediklerimizdensin,*" which means "you are the one that we were unable to make laugh," contains eight suffixes). The modern orthography of Turkish is composed of a 29-letter alphabet of eight vowels and 21 consonants, based on a modified Latin script. In most cases, a single phoneme is represented with a single grapheme, and the grapheme-to-phoneme correspondence is consistent and transparent (Raman, 1999). The only exceptions are words borrowed from other languages, which are usually transferred into Turkish with their original phonology. For example, the word "katip," which is borrowed from Arabic, is written with a single *a* but read, like in the Arabic origin, with a long *a*,/kaatip/. Syllables in Turkish (except for loan words) comprise of a single vowel. Canonical syllable structure in Turkish is CV, but other structures (such as V, VC, and CVC) also exist. Syllables with consonant clusters are rare (Raman, 1999), and the length of the vowel restricts the consonants of the coda (Kabak and Vogel, 2001). Stress position is regular and final (with some exceptions, Kabak and Vogel, 2001).

Dyslexia in Turkish

Very few studies about Turkish took the neuropsychological approach to acquired and developmental dyslexia. Or, in Raman and Weekes (2008) more positive words, "Research addressing the cognitive neuropsychology of acquired language disorders in Turkish has only just begun to flourish." The exception is a study by Raman and Weekes (2005) who studied in detail a Turkish-English bilingual stroke patient who had an acquired lexical-phonological retrieval deficit, which made him unable to read via the lexical route. This led to surface dyslexia in English, and imageability effects in reading Turkish, with good reading of non-words.

We could not find any study on types of developmental dyslexia in Turkish. The few papers that examined developmental dyslexia in Turkish did not make the distinctions between different types of developmental dyslexia and have not characterized the types of errors each of the individuals with developmental dyslexia made. Raman (2011) tested a group of students with developmental dyslexia (without identifying the type of dyslexia each of them had) and examined the effect of Age of Acquisition in word and picture naming in this group in comparison to a non-dyslexic control group. Other studies worked under the assumption that dyslexia in Turkish mainly affects reading fluency (Erden et al., 2002; Özmen, 2005; Ergül, 2012). Studies of reading in typical development focus mainly on the contribution of phonological abilities to reading and spelling acquisition both in monolingual (Öney and Durgunoğlu, 1997; Durgunoğlu, 2002; Kesikçi and Amado, 2005; Babayiğit and Stainthorp, 2010) and bilingual populations (Özdemir et al., 2012; Özata et al., 2016).

The current study, therefore, aims to start filling the gap by (1) reporting and exploring in detail a type of dyslexia

that has not been reported for Turkish yet, in either the acquired or developmental form. Exploring its characteristic error types and the properties of stimuli that are more susceptible to errors. (2) Reporting for the first time a specific type of developmental dyslexia in Turkish. (3) Examining the common belief that Turkish readers with dyslexia do not make reading errors – we will demonstrate that, when the relevant stimuli are selected, they definitely do make errors. The rich morphology and the CV structure of Turkish would further allow us to ask questions about properties of LPD that have not been tested so far: (4) What is the relative order of morphological decomposition and letter position encoding. (5) How early in the reading process are letters encoded as consonants or vowel letters.

MATERIALS AND METHODS

Participants

Dyslexic Group

The dyslexic participants in this study were 24 monolingual Turkish-speaking children in 4th grade, aged 9–10, 8 males and 16 females. All of them were right-handed. All of the participants were living in Eskişehir, Turkey. They were pupils in regular schools and regular classes. According to the reports of their parents and/or teachers, and through informal observation made by SLPs, none of them had speech and language disorders (beyond reading deficits), nor any history of brain lesion, neurological condition, or cognitive problems. None of them had been previously diagnosed as having dyslexia or learning disability. However, when we discussed their reading with their teachers, the teachers expressed concerns about the reading of almost all the participants.

To select individuals who have developmental LPD for this study, we administered the dyslexia screening task FRİGÜ (Güven and Friedmann, 2014), described in the next section. We included individuals in the LPD group according to the following inclusion criteria: significantly higher total number of errors in the screening test compared with the age-matched group, significantly more letter position errors in the screening test than the control group (and at least 6 letter position errors), and less than 10% for other types of errors.

The 24 children with LPD were identified in the following way: 16 children were identified in a school-wide reading testing in which we administered the FRİGÜ reading screening test to 299 children. The other 8 children were recruited from teachers in 6 schools (in which children from varying SES – low-middle-high– status). The teachers referred children to us who they suspected had learning or reading difficulties. We tested the reading of these children using the FRİGÜ screening test, 8 had LPD and fitted the inclusion and exclusion criteria so they were included in the study.

Control Groups

The control group for the screening task included 205 fourth graders, 111 girls and 94 boys, with no report of reading disability.

The control group for the further LPD tests included 71 fourth graders aged 9 to 10 years, 39 males and 32 females who had no speech, language, hearing or other cognitive problems, based on teacher or parent reports. Additionally, the children who were examined for the control group were tested by speech-language pathologists, who reported their clinical opinion regarding each child's language. We excluded three children who the testers suspected to have a language disorder.[1]

Procedure

Each of the participants was tested individually in a quiet room in the school. All stimuli were displayed on a white page in 14 pt. font, with double vertical spacing between words. No time limit was imposed during testing, the written word lists remained in front of the participants for as long as they needed, and no response-contingent feedback was given by the experimenter. In the silent reading tasks, we instructed the participant not to read aloud. In orally presented tests, the experimenter repeated every item as many times as the participant requested. Each of the participants took part in 13 tests, which were administered in several sessions. The number of sessions and length of each session were determined by each of the participants. The research was approved by the Ethical Committee of Anadolu University, Eskişehir, Turkey. The parents of each child signed written informed consent.

General Error Coding and Analysis

In the analysis of letter transpositions, we classified the transpositions according to the letters that participated in the transposition: consonant-consonant, vowel-consonant, and vowel-vowel migrations. If the participant produced a sequence of responses to a target word, and one of these responses was an error, we counted the item as being incorrect and analyzed the erroneous response.

Statistical Analyses

We examined whether each participant with dyslexia performed significantly below his or her age-matched control group using one-tailed Crawford and Howell's (1998) t-test. Within-participant comparisons between two conditions were conducted using chi-squared tests (two-tailed comparisons). At the group level, comparisons between two conditions were conducted using the Wilcoxon Signed-Rank test, which is the non-parametric counterpart of paired samples t-test (reported with z statistic), and more than two conditions were compared using Friedman Test. For the correlation, Spearman's rank correlation coefficient analysis was used. Effect sizes for t-tests are reported with Hedges' g, and for Wilcoxon's, when there is no normal distribution, with r (Fritz et al., 2012). Comparisons at the group level between the LPD group and the large control group were done using the t-test. An alpha level of 0.05 was used in all comparisons.

---

[1]The age-matched control group was in fact also a reading-age matched group for many of the LPD participants: if we define a reading age control group by the total rate of errors in the non-migratable words in the screening test, the 205 age-matched control participants made an average of 1.1 errors (SD = 0.9), and 13 of the LPD participants performed within the normal range for their age on these words.

THE SCREENING TEST USED TO IDENTIFY CHILDREN WITH LPD FOR THIS STUDY

The first test we administered, which we used for initial assessment and identification of individuals with LPD to include in the study, and to examine the properties of their errors, was the screening test from the FRİGÜ test battery (Güven and Friedmann, 2014), which was developed to identify types of dyslexia in Turkish. The screening part of the FRİGÜ is an oral reading test that includes three blocks: 151 single words (2–8 letters long, $M = 5.12$, $SD = 1.29$), 60 word pairs (4–5 letters long, $M = 4.88$, $SD = 0.92$), and 40 non-words (2–9 letters long, $M = 5.16$, $SD = 1.62$). The word reading block was used to identify individuals with LPD according to the criteria described above.

Some researchers have claimed that in certain languages (e.g., languages with transparent orthographies), dyslexia does not manifest itself in errors, only in slow reading. We think this is a misconception that partly results from not using the right stimuli to elicit reading errors. Our approach was, on the basis of the approach of the Tiltan reading battery (Friedmann and Gvion, 2003), to base the reading test on knowledge of the types of words and non-words that are most sensitive to each type of dyslexia, i.e., the type of stimuli in which individuals with this kind of dyslexia make most errors of the relevant type. Ever since the early days of the cognitive neuropsychological approach to dyslexia – Marshall and Newcombe (1973); Coltheart (1981), and Patterson (1981) found that dyslexias differ with respect to the types of stimuli that are most difficult in them. So there are "dimensions of words" as Patterson called it, for example, surface dyslexia is most evident in reading irregular words, phonological dyslexia in non-words, and deep dyslexia in abstract words, function words, and morphologically complex words.

The word and non-word lists of the FRİGÜ screening test were thus constructed so that they include items that are sensitive to each of the currently known types of dyslexia; words with different stress patterns or with ambiguous grapheme-to-phoneme conversion for identifying surface dyslexia; function words and morphologically complex words to identify phonological output buffer dyslexia, orthographic input buffer dyslexia, and deep dyslexia; abstract words for deep dyslexia; words (and non-words) with many orthographic neighbors for identifying orthographic analyser-output visual dyslexia, orthographic input buffer dyslexia, and letter identity dyslexia; words (and non-words) that can be read as other words by neglecting one side of the word, for identifying neglect dyslexia; and words in which vowel letter omissions, additions, migrations, or substitutions create other existing words, for the identification of vowel letter dyslexia.

The non-words were included for identifying phonological and deep dyslexia as well as various peripheral dyslexias; the word pairs were constructed such that between-word migrations create other existing words, to enable the detection of attentional dyslexia.

Importantly, the screening part of FRİGÜ was designed to also detect LPD. The list of 151 words contains 121 migratable words: 91 words in which a middle letter migration would create another existing word, and 54 words in which migration of exterior letters creates a word (24 of these words allowed for both interior and exterior migrations).

Results: Reading Screening Test

The participants with LPD made between 6 and 19 letter position errors in the single word reading block, with an average of 10.3 letter position errors ($SD = 4.1$). The control participants, on the other hand, made only 1.8 letter position errors on the average in this task (0–5 errors, $SD = 2$). Each of the LPD children performed significantly poorer than the control group [$t(204) > 1.94$, $p < 0.02$, for each of the participants].

ORAL READING OF MIGRATABLE AND NON-MIGRATABLE WORDS

Now that the screening task identified 24 children who had LPD, we continued with a line of tests that were developed to examine the nature of this dyslexia, and the way it is manifested in Turkish. We created a list of 183 migratable words to allow for the in-depth assessment of the effect of morphology on migrations (see section "Does Morphological Analysis Precede Letter Position Encoding? Assessing the Interaction of Morphology and Migrations"); the effect of the consonant-vowel status on migration (see section "Is Letter Position Encoding Sensitive to the Consonant-Vowel Status of the Target Letters?"); the position of the migrating letters within the word (assessing middle-exterior, adjacent-non-adjacent, within-across syllable, and length effect, see section "Further Analyses of the Properties of Letter Migrations in Turkish"); and frequency effect (see section "What Is the Nature of the Letter Position Encoding Deficit? Incorrect Underspecified Encoding? The Effect of Frequency on Migrations and Its Theoretical Implication").

Experimental Stimuli

The migratable word list included 183 migratable words, 4-to-8 letters long ($M = 5.2$, $SD = 0.9$). Each of these words was such that at least one letter migration within the word results in an existing word (see examples for various types of migrations in **Table 1**). Each of the words in the list was also such that a letter identity error could create another existing word.

Results

Migrations in Reading Migratable Words

The LPD participants made a total of 433 migration errors in reading the migratable word list. **Figure 1** summarizes the letter position error rates the children with LPD made. Each of the 24 children with LPD made significantly more migrations than the control group ($p < 0.001$, using Crawford and Howell's, 1998, t-test for the comparison of an individual to a control group). The difference was also significant at the group level, where the LPD group made significantly more letter migrations (10% migrations) than the control group (who made only 1%

TABLE 1 | Examples for migration errors of various types that the LPD participants made.

Condition	Target word	Response with migration error	Translation word	Translation response	Example from English
Middle migrations					
Adjacent consonant-consonant (CC) migration	altı	atlı	six	horseman	badly-baldy
Adjacent consonant-vowel (CV) migration	alınma	alnıma	to take offense	to my forehead	from-form
Consonant-consonant migration across a vowel (C-C)	ebedi	edebi	eternal	literary	slime-smile
Vowel-vowel migration across a consonant (V-V)	çelik	çilek	steel	strawberry	toner-tenor
Exterior migrations[a]					
Adjacent consonant-vowel (CV) migration	atkı	takı	scarf	jewel	acres-cares
Consonant-consonant migration across a vowel (C-C)	yakın	yanık	close	burn	inlet-intel
Vowel-vowel (Exterior) migration across a consonant (V-V)	katı	kıta	stiff	continent	demo-dome

[a]*There were no words with a potential of exterior adjacent CC migrations, due to the Turkish syllable structure.*
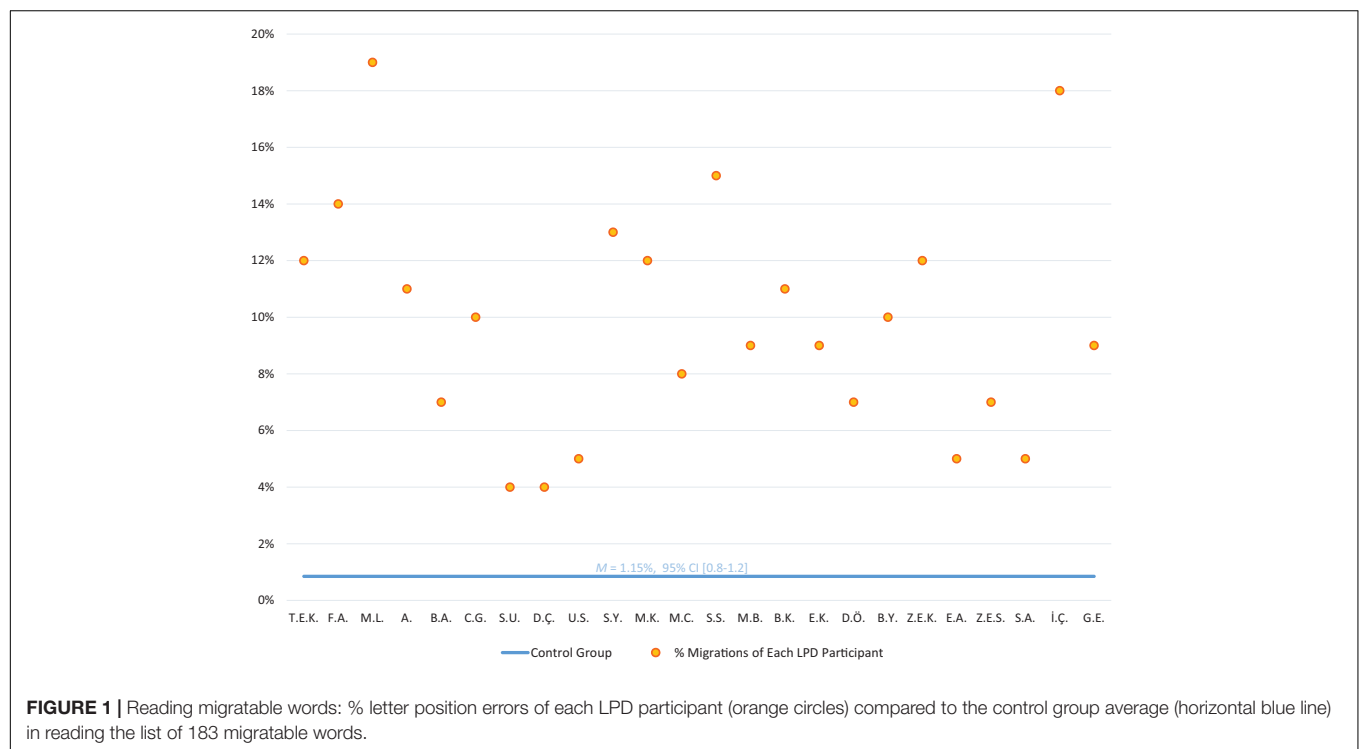
FIGURE 1 | Reading migratable words: % letter position errors of each LPD participant (orange circles) compared to the control group average (horizontal blue line) in reading the list of 183 migratable words.

migrations on this list, $SD = 1.27$), $t(24) = 10.71$, $p < 0.0001$, with a very large effect size ($g = 2.53$). See **Table 1** for examples for the various types of migrations that the participants with LPD made.

Refuting a Misconception: Turkish Dyslexics Do Make Errors When the Appropriate Stimuli Are Presented: Migratable vs. Non-migratable Words

We asked whether the participants make more migration errors when the target words are migratable, i.e., words in which a migration error can create another existing word (like the English word *form,* in which a migration error can create *from*) than on non-migratable words.

We compared the migrations in the list of 183 migratable words to the rate of migrations in reading a list of non-migratable words, which included words in which no single transposition created an existing word (e.g., the target word *gözlük* [glasses], for which all possible migrations result in non-lexical responses). This non-migratable word list included 32 words, 4-to-7 letters long ($M = 4.9$, $SD = 0.7$), with a relatively high frequency (all were among the 5000 most frequent words in Turkish, and more than half of them were among the top 2000 most frequent words, according to Aksan et al. (2016) frequency data.

This analysis showed a striking difference between migratable and non-migratable words: the children with LPD made

TABLE 2 | % migration errors in migratable and non-migratable words.

Participant	Migratable words (*N* = 183)	Non-migratable words (*N* = 32)
T.E.K.	12	0
F.A.	14	0
M.L.	19	0
A.C.	11	0
B.A.	7	0
C.G.	10	0
S.U.	4	0
D.Ç.	4	0
U.S.	5	0
S.Y.	13	0
M.K.	12	0
M.C.	8	0
S.S.	15	0
M.B.	9	0
B.K.	11	0
E.K.	9	0
D.Ö.	7	0
B.Y.	10	0
Z.E.K.	12	0
E.A.	5	0
Z.E.S.	7	0
S.A.	5	0
İ.Ç.	18	0
G.E.	9	0
Average (*SD*)	**10 (4)**	**0**

N = number of words that allow for at least one error of the relevant type.

no migrations in non-migratable words, whereas they made an average of 10% migrations in reading the migratable words (**Table 2**).

Relatedly, to examine whether the letter position errors of the participants with LPD tended to create existing words, we analyzed their migration responses in reading migratable words. For each migratable word, there was at least one option for a migration that yields an existing word, and at least one other option for migration response that yields a non-word. We examined whether the participants' migrations tended to create existing words.

This analysis showed that when they read a migratable word with a migration error, most of their responses were lexical (76.6% of their migration responses were lexical, *SD* = 31.2%). The same was true in reading migratable non-words, in which, as we report below in Section "What Is the Locus of LPD in the Reading Model: Nonword Reading and Silent Reading Tasks – Results", most errors were lexical as well.

The tendency to produce lexical responses guided us in the analyses of the characteristics of the participants' migrations: for each analysis, we calculated the rate of errors out of the number of target words in which such errors would create an existing word, i.e., words that have a lexical potential for the relevant error type (for example, we calculated adjacent migrations not out of all words, but rather only out of the target words in which a migration of adjacent letters creates an existing word).

DOES MORPHOLOGICAL ANALYSIS PRECEDE LETTER POSITION ENCODING? ASSESSING THE INTERACTION OF MORPHOLOGY AND MIGRATIONS

Turkish has a very rich morphology, so studying LPD in Turkish readers allowed us to examine a theoretical question about the interaction between letter position encoding and morphological analysis. Specifically, we were interested in the relative order in which letter position encoding and morphological analysis take place. If morphological analysis precedes letter position encoding, then the morphological structure of the target word should affect letter position errors, and migrations should occur only within a morpheme. If, however, letter position encoding precedes morphological analysis, then the morphological structure of the target word should not affect migration errors.

To examine this question, we used three kinds of analysis. We examined whether morphologically complex words yielded a different rate of letter position errors than monomorphemic words. We also tested whether letter migrations changed the morphological structure of the target word and whether migrations occurred only within morpheme or also across morphemes.

Analyses and Results

The first analysis examined whether morphologically complex words yield a different rate of letter position errors than morphologically simple words. For this analysis, we compared 44 morphologically-complex words from the migratable word list (which includes words with derivational and words with inflectional morphemes) to 78 monomorphemic migratable words. We selected these words so that they would be matched on length, which led us to include 4–6 letters long morphologically complex words (*M* = 5.39, *SD* = 0.61) and 5–7 letters long monomorphemic words (*M* = 5.22, *SD* = 0.47), so the word lengths in the two groups did not differ significantly.

The results were such that the children with LPD had very similar rates of migrations on the morphologically complex words (8.9%) and on the morphologically simple words (9.5%), and this difference was not significant (*Wilcoxon z* = 0.24, *p* = 0.81).

In the second analysis, we examined whether the migrations changed the morphological structure of the target word. This analysis indicated that there were quite a few migrations that changed the morphological structure of the target word: 21% of the migrations changed the morphological structure of the target word (*SD* = 19%). For example, the monomorphemic target word *akran* (peer) was read with a transposition as "arkan", a morphologically complex word, constructed from *arka* (back), and the suffix –*n* (singular 2nd person possessor).

Or the morphologically simple target word *eskiz* (sketch), which was read with a transposition as the morphologically complex "eksiz" in which the stem *ek* means "supplement" (or, to confuse us, "a morpheme"), and the suffix *siz* means "without" (so this morphologically complex word actually means "without a morpheme" i.e., a monomorphemic word).

The final analysis tested whether migrations occurred only within-morpheme or also across morphemes. This analysis indicated that, in reading the morphologically complex words, the participants transposed letters of the stem with letters of the non-stem morpheme (e.g., *konulu*: konu-lu, themed →*kolunu*: kolu-nu, your arm). Such cross-morpheme migrations occurred on 6.3% of the morphologically complex words ($SD = 3.7$). In fact, migrations within the stem, which involved two letters of the stem (*yenile*: yeni-le, renew → *yinele*: yine-le, repeat) and migrations within the non-stem morpheme (*kirli*: kir-li, dirty → *kiril*, Cyrillic) occurred less frequently than across-morpheme migrations (within the stem: 2% of the morphologically complex words, $SD = 2.3$, within the non-stem morpheme: 0.7%, $SD = 1.1$).

These three kinds of evidence for the lack of sensitivity of letter position errors to the morphological structure of the target word suggest that LPD affects a stage that precedes morphological analysis, and hence, it is not sensitive to the morphological structure of the target words.

IS LETTER POSITION ENCODING SENSITIVE TO THE CONSONANT-VOWEL STATUS OF THE TARGET LETTERS?

To examine the theoretical question of when the consonant-vowel distinction becomes accessible during the process of single word reading, we examined the effect of the consonant-vowel status of a letter on the rate of migrations.

We asked two main questions: whether consonant and vowel letters are differentially susceptible to migrations, and whether they tend to migrate more within their class (consonants transpose with consonants and vowels with vowels) than across class.

Analyses

For this sake, we compared four types of migration:

(1) V_1CV_2: a transposition of two vowels across a consonant [e.g., *eli* (his/her hand) > "ile" (with)]
(2) C_1VC_2: a transposition of two consonants across a vowel [e.g., *tak* (attach) > "kat" (floor)]
(3) CC: a transposition of two adjacent consonants: [e.g., *atlı* (horseman) > "altı" (six)]
(4) CV: a transposition of adjacent consonant and vowel [e.g., *atlı* (horseman) > "atıl" (idle)]

We did this analysis only for the participants who did not have vowel dyslexia (Khentov-Kraus and Friedmann, 2018) in

addition to LPD, because a higher rate of migrations that involve vowels may, in their case, be a result of their vowel dyslexia.

Within the 183 migratable words list, there were 98 words in which a consonant-consonant migration creates another word (51 words allowing for Adjacent CC migrations and 67 allowing for CC migration across a vowel); 63 words that allow for adjacent consonant-vowel migration, and 74 words that allow for a vowel-vowel transposition across a consonant.

Results

The results, summarized in **Table 3**, indicated a clear difference between the different kinds of migration. Migrations that involved only consonants (either adjacent CC transposition or CVC- transposition of two consonants across a vowel) were the most common type of migration (14%), whereas migrations that involved only vowel letters, or migrations that involved a consonant and a vowel occurred less often (5% each).

TABLE 3 | Consonant and vowel letter migrations of the participants with LPD in migratable word reading (% migrations of each type out of the number of migratable words with a lexical potential for such error).

Participant	All C-C (N = 98)	C-C across V (N = 67)	V-V across C (N = 74)	Adjacent CC (N = 51)	Adjacent CV (N = 63)
F.A.	23	16	3	18	6
M.L.	30	18	1	22	17
A.C.	20	13	0	16	5
B.A.	8	4	5	6	5
C.G.	19	13	0	18	2
S.U.	7	4	1	6	2
D.Ç.	5	3	0	4	5
U.S.	8	6	3	8	0
S.Y.	13	10	3	10	14
M.K.	21	22	1	10	2
M.C.	12	12	4	8	0
M.B.	12	6	1	14	8
B.K.	16	10	3	18	3
E.K.	11	6	4	12	5
D.Ö.	6	3	0	4	14
B.Y.	13	6	5	18	3
Z.E.K.	16	12	4	16	5
E.A.	7	3	1	8	3
Z.E.S.	11	3	1	16	2
S.A.	6	3	1	12	3
İ.Ç.	26	28	0	12	13
G.E.	9	9	9	6	0
Average (SD)	**14(7)**	**10(7)**	**2 (2)**	**12(5)**	**5(5)**

Participants with LPD and vowel dyslexia

S.S.	9	4	12	12	14
T.E.K.	11	7	12	2	11

C = consonant letter, V = vowel letter, N = number of words in which at least one error of the relevant type creates an existing word (some words allow for more than one type of error: e.g., the target word "istem" can be read with adjacent CC error as itsem, or with a C-C error across V as ismet. So it is counted once as a word with any C-C migration potential, and once in words with adjacent CC migrations potential, and also once in words with CC across V migrations potential). Due to the syllable structure in Turkish, we had no words with adjacent VV migration potential and no non-adjacent CV.

A Friedman's test indicated that the difference between the three types of migration (C with C, C with V, and V with V) was significant, $\chi^2(2) = 30.45$, $p < 0.001$, with consonant-only migrations being the most frequent migrations.

Consonants migrated across vowels ($C_1 V C_2 \rightarrow C_2 V C_1$) significantly more frequently than vowels across a consonant ($V_1 C V_2 \rightarrow V_2 C V_1$), *Wilcoxon z* = 4.42, $p < 0.0001$, $r = 0.64$. Consonants transposed with an adjacent consonant (CC) significantly more often than with adjacent vowels (CV), *Wilcoxon z* = 3.85, $p < 0.0001$, $r = 0.56$.

Another type of analysis that points in exactly the same direction is the analysis of the "preferred migration type" in target words that allow for several types of migrations (C-C, V-V, V-C). There were 64 such target words, and the participants showed the same tendency toward consonant migrations: when they read a word in which several types of migrations were possible, they most often made a C-C migration (67% of the migrations on these words) and had much fewer V-V migrations (17%) or C-V migrations (16%). The C-C migrations were significantly more frequent than the other migration types, Friedman's test $\chi^2 = 21.62$, $p < 0.001$.

The results show clearly that consonant letters are more susceptible to transpositions than vowel letters, and that consonant-only transpositions occur more frequently than transpositions that involve a consonant and a vowel letter. Beyond its bearing on the characterization of LPD, this finding that indicates that the classification of letters to consonants and vowels happens very early in the process of orthographic-visual analysis, before or together with letter position encoding, and that consonant letters are processed separately and differently from vowel letters.[2] It is also interesting to note that the two children who had vowel dyslexia in addition to letter position dyslexia (SS and TEK, presented in the bottom of **Table 3**) made more transpositions that involve a vowel than transpositions that include only consonants.

FURTHER ANALYSES OF THE PROPERTIES OF LETTER MIGRATIONS IN TURKISH

Middle vs. Exterior Letter Migrations

Studies of LPD in Hebrew, Arabic, and English report that individuals with LPD make more migrations that involved only middle letters than migrations that involve an exterior letter (Friedmann and Gvion, 2001; Friedmann and Rahamim, 2007; Friedmann and Haddad-Hanna, 2012, 2014; Kohnen et al., 2012).

To examine whether this was also the case for Turkish LPD, we compared the rates of migrations that involved only middle letters and migrations that also involved an exterior (first or final) letter.

The 183 migratable words list included 91 words that have at least one possibility for middle migration, 34 words with a possibility for a migration that involves an exterior letter, and 58 words with a potential for both middle and exterior migrations.

The results, presented in **Table 4**, show that the Turkish-readers with LPD made both middle letter migrations and exterior letter migrations, but they, like LPD participants in the other languages tested, made significantly more migrations of middle letters (9%) than migrations that involved an exterior letter (6%), *Wilcoxon z* = 2.73, $p = 0.006$, $r = 0.39$.

This predominance of middle migrations can also be seen in another type of analysis that assesses the "preferred migration type" in target words in which both middle and exterior migrations create other existing words. There were 55 such target words, and the participants showed the same tendency toward middle migrations: when they read a word in which both a middle migration and an exterior migration would create existing words, they made almost three times more middle migrations (73%) than exterior ones (27%), a difference that was significant, *Wilcoxon* z = 5.06 $p < 0.0001$.

Migrations of Adjacent and Non-adjacent Letters

Hebrew readers with LPD make more migrations in adjacent letters than non-adjacent letters. We tested whether this is the case also for Turkish children with LPD.

Within the 183 migratable words list, in 61 words only adjacent letter migrations created other existing words, 95 words allowed only for non-adjacent letter migration, and 27 words had a lexical potential for both adjacent and non-adjacent letter migration.

The results, summarized in **Table 4**, show that the Turkish LPD participants made significantly more migrations in adjacent letters (12%) than in non-adjacent letters (7%), *Wilcoxon z* = 3.09, $p = 0.002$, $r = 0.45$. However, as we report in Section "Is Letter Position Encoding Sensitive to the Consonant-Vowel Status of the Target Letters?," the consonant-vowel status of the letter affects migration considerably; once the consonant-vowel status is kept constant (analyzing only consonant-consonant migrations), the size of the difference between adjacent and non-adjacent letters shrinks, *Wilcoxon z* = 2.02, $p = 0.04$.

Migrations Across and Within Syllables

We also investigated whether more migrations occur within- or across syllables. The results, presented in **Table 4**, show that there were significantly more across-syllable migrations (8%) than within-syllable migrations (5%), *Wilcoxon z* = 3.07, $p = 0.002$, $r = 0.44$.

Length Effect

To examine the effect of word length on the rate of migrations, we analyzed the participants' migration rates in reading aloud

---

[2]The difference between consonant-only migrations and migrations that involved a vowel letter is not due to frequency differences between the conditions, which were evenly distributed between the relative-frequency conditions. Within each frequency condition – similar frequency between target and migration counterpart, target more frequent than migration, migration result more frequent than target – the LPD participants made far more migrations that involved only consonant letters than migrations that involved vowel letters (CV and VV) 25% CC and 9% CV/VV in the frequent migration counterpart condition, 13% CC and 6% CV/VV in the similar frequency condition, and 12% CC and 3% CV/VV in the frequent target word condition.

TABLE 4 | Percentages migrations of the various kinds that the participants with LPD made in reading migratable words.

Participant	Total (N = 183)	Adjacent (N = 86)	Non-adjacent (N = 120)	Middle (N = 149)	Exterior (N = 90)	Across syllable (N = 153)	Within syllable (N = 94)
T.E.K.	12	10	11	10	8	10	6
F.A.	14	16	10	13	7	14	4
M.L.	19	23	13	15	13	17	10
A.C	11	13	8	10	6	11	3
B.A.	7	7	6	6	4	6	4
C.G.	10	10	8	9	6	9	5
S.U.	4	5	3	5	1	4	2
D.Ç.	4	5	3	2	4	3	3
U.S.	5	5	5	5	2	5	3
S.Y.	13	16	8	11	7	11	6
M.K.	12	7	13	7	12	11	5
M.C.	8	5	9	9	2	8	3
S.S.	15	17	10	12	10	10	12
M.B.	9	14	4	7	7	7	6
B.K.	11	13	8	10	6	8	7
E.K.	9	10	6	9	3	7	6
D.Ö.	7	13	2	4	8	4	7
B.Y.	10	14	6	11	2	10	4
Z.E.K.	12	14	8	13	2	10	6
E.A.	5	7	3	4	3	3	4
Z.E.S.	7	12	2	7	1	7	2
S.A.	5	7	3	5	1	5	2
İ.Ç.	18	17	15	13	16	10	19
G.E.	9	16	2	9	3	9	2
Average (SD)	**10 (4)**	**12 (5)**	**7 (4)**	**9 (3)**	**6 (4)**	**8 (4)**	**5 (4)**

N = number of words that allow for at least one error of the relevant type.

TABLE 5 | The relationship between the number of letters and LPD error rate.

Word Length	Number of words	% migrations
4 letters	31	10
5 letters	101	9
6 letters	33	8
7 letters[a]	18	19

[a] This group included 17 7-letter words and a single 8-letter word.

the list of 183 4–8 letter migratable words (see section "Experimental Stimuli").

We conducted Spearman's Rho correlation coefficient analysis to see if there is a relationship between length and migration errors. A correlation analysis of the results (presented in **Table 5**) indicated that the correlation coefficient was low and non-significant ($R = -0.11$, $p = 0.19$).

WHAT IS THE LOCUS OF LPD IN THE READING MODEL: NON-WORD READING AND SILENT READING TASKS

The next question was where in the word-reading model is the impairment that gives rise to LPD. For this sake, we examine

the participants' non-word reading and compare it to their word reading (section "Non-word Reading" below): if they make migration errors in non-words as well, this would mean that the deficit is not in lexical components and that it is rather in a component that is shared by lexical and non-lexical processes. We then test the participants' silent reading using various tasks to examine whether the locus is in the orthographic-visual analysis stage. If it is, a deficit in the orthographic-visual analysis should cause migrations not only in reading aloud but also in lexical decision and comprehension of written words (section "Silent Reading Tasks" below). If the deficit is indeed in the orthographic input, phonological output should not show migrations when the input does not involve reading. This we tested in Section "Assessing Phonological Output Using Non-word and Word Repetition," which examined the participants' phonological output using non-word and word repetition.

Non-word Reading

To examine how the LPD participants read non-words, and to further test whether their deficit was pre-lexical or at a lexical stage, we presented them with an additional list of 60 non-words. The non-words in the list were 4-to-6 letters long ($M = 4.92$, $SD = 0.42$). Half of the non-words (30) were migratable, i.e., non-words in which a letter migration creates an existing word (e.g., the non-word "bakrı" is migratable because the migration

of the two middle consonants creates the word "barkı," his/her home). The other 30 non-words were non-migratable so that no migration created another word (e.g., "solik" or "bikeş").

The 30 migratable non-words were selected to allow the examination of the characteristics of migrations also in non-words. To compare middle and exterior letter migrations, 27 of the non-words were such that at least one migration of an exterior letter would create an existing word, and 10 non-words were such that migration of middle letters would create an existing word (7 of these words had both middle and exterior migration potential). There were 15 words that had a potential of consonant-consonant transposition, 3 words with a potential of vowel-vowel transposition, and 25 words with vowel-consonant transposition potential.

Results

The 24 participants with LPD made an average of 18% migrations when reading the migratable non-words. The rate is significantly higher than that of the control group (2%), $t(25) = 7.63$, $p < 0.0001$, $g = 1.8$. On the individual level (see **Figure 2**), 21 children with LPD made significantly more migration errors than the control group (14 children $p < 0.001$, and 7 children $p < 0.05$, using Crawford and Howell's (1998), t-test).

The participants made significantly more migrations in reading the migratable non-words (18%) than in reading the non-migratable non-words, where they made 7% migrations (Wilcoxon $z = 4.41$, $p = 0.0001$, $r = 0.64$). Still, the rate of migrations that the children with LPD made in non-migratable non-words was significantly larger than that of the control group, who only made 0.3% migrations in non-migratable non-words, $t(23) = 4.92$, $p < 0.0001$, $g = 1.16$. This finding is important because it indicates that the deficit that underlies the migration errors in LPD is not in the orthographic input lexicon but

rather in an earlier stage that affects words, non-words, and non-migratable non-words: the orthographic-visual analyzer.

When we analyzed the errors that the LPD participants made in reading the migratable non-words (presented in detail in **Table 6**), we see that, like in their reading of the existing migratable words, they made significantly more migrations in middle letters (31%) than migrations that involved an exterior letter (9%; Wilcoxon $z = 4.51$, $p = 0.0001$, $r = 0.65$). They made significantly more migrations that involved only consonants (22%) than migrations of a consonant and a vowel (11%), Wilcoxon $z = 3.01$, $p = 0.001$, $r = 0.44$, and significantly more migrations across syllables (24%) than within a syllable (9%), Wilcoxon $z = 4.30$, $p = 0.0001$, $r = 0.62$.

Like in the word reading, most of the error responses the LPD participants made in reading the migratable non-words were lexical ($M = 81.3\%$ of the total errors in migratable non-words, $SD = 17.1\%$), with significantly more lexical errors than non-lexical errors, Wilcoxon $z = 5.36$, $p < 0.001$, $r = 0.77$. When we only look at migration responses in reading the migratable non-word list, the picture remains the same: most of their migration responses were lexical ($M = 88.6\%$ of their transposition errors in migratable non-words, $SD = 44.3\%$; significantly more lexical than non-lexical migration responses, Wilcoxon $z = 5.63$, $p < 0.001$, $r = 0.82$).

Not surprisingly, when they made migration errors in reading non-migratable non-words, where migrations could not yield an existing word, they produced mainly non-lexical responses ($M = 33.6\%$ of the total errors in non-migratable non-words, $SD = 35.1\%$, Wilcoxon $z = 2.87$, $p = 0.002$, $r = 0.42$).

Silent Reading Tasks

If the source of letter migrations is indeed in a deficit in the orthographic-visual analysis stage, we would expect migrations
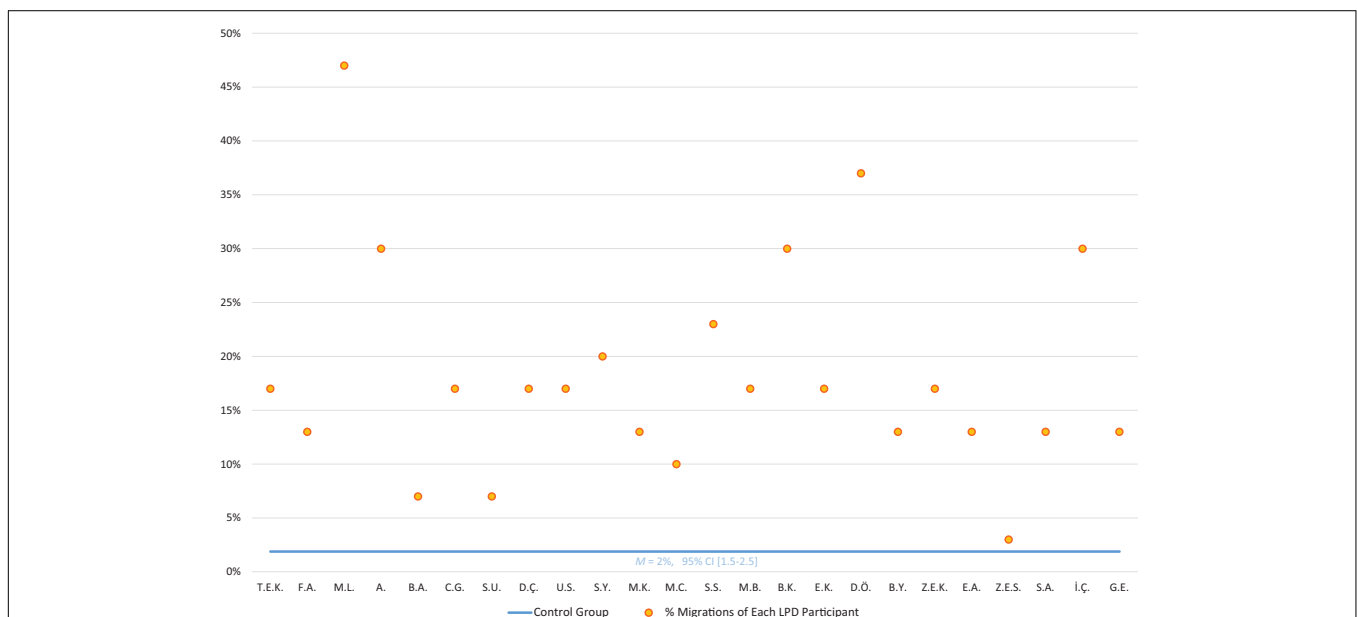

FIGURE 2 | Migratable non-word reading: % letter position errors of each LPD participant orange circles) compared to the control group (horizontal blue line).

TABLE 6 | Characteristics of migration errors in reading migratable non-words (% errors in each of the stimulus types).

Participant	Total (N = 30)	Middle (N = 10)	Exterior (N = 27)	Across syllable (N = 14)	Within syllable (N = 25)	C-C (N = 15)	V-V (N = 3)	C-V (25)	Adj. CC (N = 13)	Adj. CV (N = 19)
T.E.K.	17	10	15	7	16	13	0	16	8	16
F.A.	13	30	4	21	4	20	0	4	23	5
M.L.	47	50	33	36	36	33	0	44	31	47
A.C.	30	60	11	57	4	40	33	12	46	11
B.A.	7	10	4	14	0	7	33	4	8	0
C.G.	17	50	0	36	0	33	0	0	38	0
S.U.	7	20	0	14	0	13	0	0	15	0
D.Ç.	17	50	0	36	0	33	0	0	38	0
U.S.	17	30	7	29	4	13	33	12	15	11
S.Y.	20	30	11	21	12	27	0	12	23	11
M.K.	13	30	4	21	4	27	0	4	23	0
M.C.	10	10	7	14	4	13	33	8	8	0
S.S.	23	40	7	29	12	27	0	16	23	16
M.B.	17	30	7	21	8	27	0	8	31	0
B.K.	30	50	15	43	12	33	67	20	31	11
E.K.	17	40	4	21	8	20	0	8	31	5
D.Ö.	37	30	30	29	28	20	33	32	31	32
B.Y.	13	10	11	7	12	7	0	12	8	16
Z.E.K.	17	10	15	7	16	7	0	16	8	21
E.A.	13	40	0	29	0	27	0	0	31	0
Z.E.S.	3	10	0	7	0	7	0	0	8	0
S.A.	13	20	7	21	4	13	33	8	15	5
İ.Ç.	33	50	19	43	16	47	0	20	38	11
G.E.	13	30	4	21	4	27	0	4	23	0
Average (SD)	**18 (10)**	**31 (16)**	**9 (9)**	**24 (13)**	**9 (9)**	**22 (11)**	**11 (19)**	**11 (11)**	**23 (12)**	**9 (12)**

C = consonant letter, V = vowel letter, N = number of words that allow for at least one error of the relevant type, Adj. = Adjacent.

to occur in silent reading tasks that do not involve reading aloud. To examine this, we ran 3 reading tasks that did not involve reading aloud: lexical decision, same-different decision, and comprehension.

Lexical Decision

The word list for lexical decision included 59 items: 29 words and 30 non-words, 4–5 letters long ($M = 4.89$, $SD = 0.37$). All the words and non-words were migratable. We asked the participants to read the list silently and mark only the existing words.

Results

The participants with LPD made an average of 19% ($SD = 12.9\%$) errors on the lexical decision task, significantly more than the control group (8%, $SD = 5.4\%$), $t(25) = 4.06$, $p = 0.0004$, $g = 0.96$. The participants made 20% errors of accepting migratable non-words as existing words, and 18% errors of judging existing words as non-words. The individual performance of the LPD participants is presented in **Figure 3**. In the individual level analysis, 16 of the 24 LPD participants performed significantly poorer than the control group ($p < 0.05$).

Same-Different Decision

In the same-different task, the participants were presented with 60 written pairs of 4–7 letter words ($M = 5.08$ letters, $SD = 0.67$), presented side by side with a single space between them. Half of

the pairs (30 pairs) included two migratable words that differed in the position of the middle letters. The other 30 pairs included identical migratable words. We asked the participants to decide, for each pair, whether the two words were the same or different.

Results

The participants with LPD made significantly more errors in this task (9%) than the control group (who made 2% errors), $t(23) = 2.83$, $p = 0.009$, $g = 0.67$. The LPD participants made 9% errors in which they said "same" for pairs of words that differed in the order of letters, and 10% errors in which they said "different" for identical pairs of migratable words. The analysis of the individual performance of the LPD participants, presented in **Figure 4**, showed that 12 of the 24 participants performed significantly poorer than the control group ($p < 0.05$).

Comprehension Task: Migratable Word Association

We assessed the comprehension of migratable words using a word association task. The task included 28 items. Each item included 4 words: a target migratable word and 3 words from which the participant needed to select one. The target migratable word allowed for at least two different migrations that can create existing words. The three options included one word that is semantically related to the target word (e.g., for the target word *eksi*, minus, the semantically related word was *negatif*, negative). The two distractor words were semantically related to possible
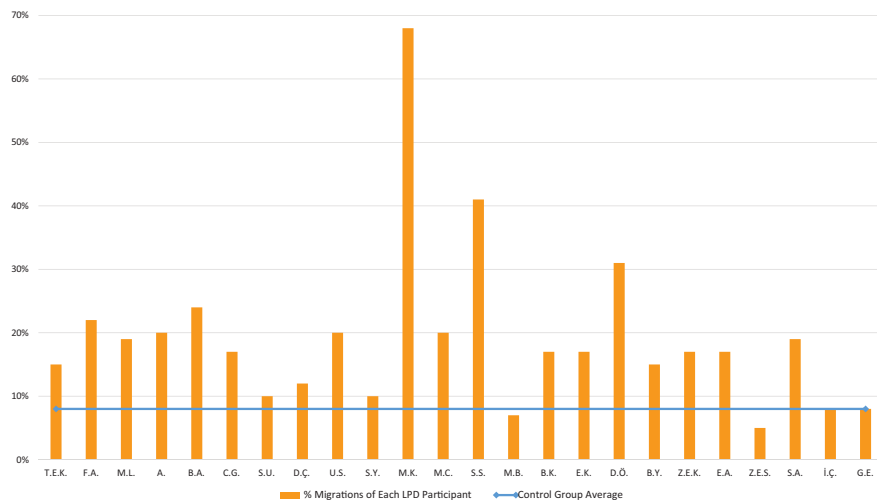
FIGURE 3 | Lexical decision: % letter position errors of each LPD participant (orange columns) compared to the control group average (blue horizontal line).
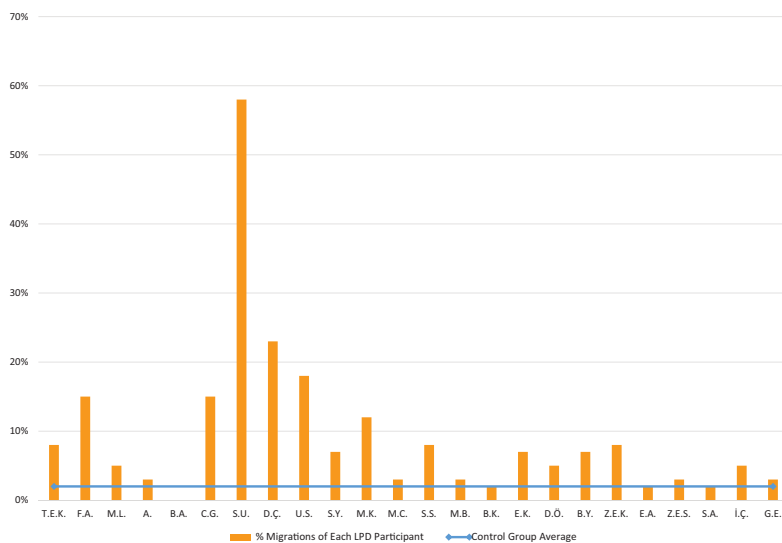

FIGURE 4 | Same-different task: % letter position errors of each LPD participant (orange columns) compared to the control group average (blue horizontal line).

migration counterparts of the target word. For example, "eksi" can be read with migration as "eski" (old) and as "kesi" (cut), so for the first migration, we have presented the word "yeni" (new), and for the second migration "bıçak" (knife). We selected target words according to the characteristics that we knew induced more migrations in our participants' reading: most of the target words had a potential for middle CC migration, and the target words were less frequent than their migration result, which were semantically related to the distractors. We tried as much as possible to use non-migratable words for the three options (88% of the options were non-migratable).

The target word was presented in orange on the left, and the three options were presented in black, one above the other to its right, in random order. We requested the participants to select the word that was most related to the target word.

Results

The children with LPD made 33% errors in this task, an error rate that was significantly higher than that of the control group (which was only 5%), $t(24) = 9.61$, $p < 0.0001$, $g = 2.27$. Each of the LPD participants made significantly more errors in this task than the control group (for 20 LPD children, $p < 0.001$; for the rest 4 children, $p < 0.05$), see **Figure 5** for the performance of each participant.

Assessing Phonological Output Using Non-word and Word Repetition

In order to further explore the locus of impairment that gives rise to LPD and to examine an alternative explanation according to which the migration errors resulted from a deficit
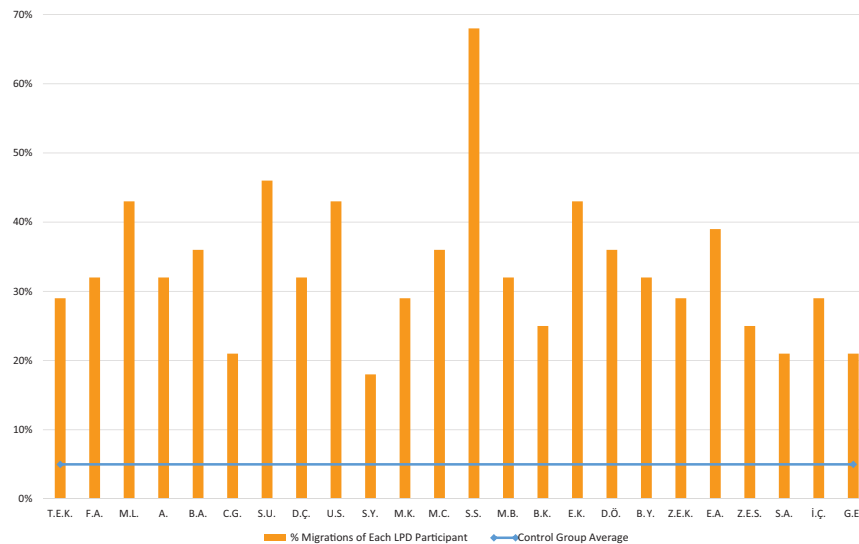
FIGURE 5 | Comprehension task: % letter position errors of each LPD participant (orange columns) compared to the control group average (blue horizontal line).

in the phonological output buffer, we assessed these children's phonological output using a non-word repetition task and a task of repeating words that the participant had read with a migration error.

Non-word Repetition

The participants repeated non-words using a standardized non-word repetition task (Turkish Non-word Repetition Test, Topbaş et al., 2014). The test included 30 items (1–5 syllables long), which consisted of 15 non-words that violate Turkish phonotactic constraints, 10 non-words that obey Turkish phonotactic rules, and 5 morphologically complex non-words. The test is normed with 150 typically developing children.

Results

All of the 16 LPD children who participated in the non-word repetition task performed this task within the normal range, with scores above the threshold for impaired repetition. The mean number of correct repetition of the LPD participants was 27.7 (out of 30), $SD = 1.47$.

Migratable Word Repetition

For each of the 19 children who participated in this task, we selected 10 of the migratable words that they read with a migration error and we then asked them to repeat these words.

Results

All of the 19 LPD children who participated in the migratable word repetition task performed this task flawlessly, with no migration error, and in fact, with no other error too.

The results of the two repetition tasks indicate that the participants had no phonological output buffer deficit, and support our conclusion, reached on the basis of the silent reading tasks, that the origin of the deficit that underlies LPD is in the input reading stages.

Theoretical Conclusion: LPD Is a Deficit in the Letter Position Encoding Function in the Orthographic-Visual Analysis Stage

The results of the three silent reading tasks: same-different decision, lexical decision, and written word comprehension (summarized in **Figure 6**) all point to the same conclusion: LPD affects not only reading aloud but also tasks that involve reading without oral production. These results, together with the findings that LPD affected both words and non-words, point to the locus of impairment in the reading model as a deficit that affects the early pre-lexical stage of orthographic-visual analysis rather than the phonological output stages. This conclusion is supported by the normal phonological output abilities the participants demonstrated in non-word and migratable words repetition.

WHAT IS THE NATURE OF THE LETTER POSITION ENCODING DEFICIT? INCORRECT OR UNDERSPECIFIED ENCODING? THE EFFECT OF FREQUENCY ON MIGRATIONS AND ITS THEORETICAL IMPLICATION

Examining the effect of the relative frequency of the target word and its migration counterpart can shed light on the nature of the letter position encoding deficit in LPD. Two options are imaginable: incorrect position encoding and underspecified position. If the nature of the letter position deficit is incorrect encoding of letter positions, the word with the incorrect positions is identified in the orthographic input lexicon according to the incorrect information that arrived from the orthographic-visual
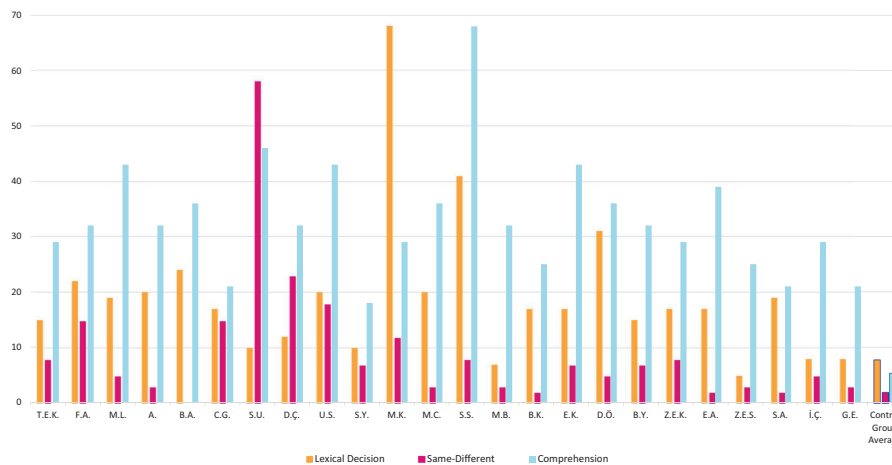
FIGURE 6 | Percentage errors of each of the LPD participants on the three silent reading tasks in comparison to the control group (control data reported in the rightmost columns).

analyzer and no effect of frequency is expected. If, however, the nature of the letter position encoding deficit is that the position of some (usually middle) letters is not encoded, frequency should affect the error rates. This is because letter position is encoded at a stage before the orthographic input lexicon, and partial position information that is transferred to the orthographic lexicon, which is organized by word frequency, should first activate the more frequent word of the migratable word pair. Thus, according to the partial letter position encoding hypothesis, when the target word is less frequent than its migration counterpart, the less frequent word is expected to be read as the more frequent one; the more frequent target word is expected to be read with fewer migrations.

Method

In order to examine the effect of frequency on letter migrations in children with LPD, we wanted to use a frequency rating that is appropriate for their age and their familiar world. We, therefore, collected frequency ratings from 30 typically-developing children in the same age and classes as our LPD participants. We presented them with 305 pairs of migratable word pairs – the target word we used in the test and its possible migration result. We asked the children to judge, for each pair, which of the two words was more familiar to them, and occurred more frequently in what they read. We collected their judgments and defined, for each pair, which word was more frequent. Then, we selected the target-response pairs for which there was a clear frequency difference.[3]

Results

The results indicated that there were far more migrations when the target word was clearly less frequent than the migration result (21.2% migrations on the average) than when the target

word was clearly more frequent than its migration result (9.0% migrations on average). This comparison was significant (*Wilcoxon z* = 4.24, $p < 0.0001$, $r = 0.61$).[4] This indicates that frequency affects migrations and supports the partial position encoding hypothesis.

NUMBER READING

To examine whether LPD results from a general visual/perceptual deficit or whether it rather pertains to orthographic material only, we examined the LPD children's reading of multi-digit numbers. We presented 40 multi-digit numbers (2–4 digits long, $M = 3.1$ digits, $SD = 0.8$), and asked the participants to read each number aloud.

Results

The LPD participants, who made a considerable rate of migrations in reading words, made very few migration errors when they read numbers. They made only 0.25% migration errors in reading multi-digit numbers aloud. Of the 24 children, 21 made no digit migrations at all in reading numbers, and three children made a single digit migrations error. This migration rate in numbers was significantly smaller than the rate of migrations that the same children made in reading words, *Wilcoxon z* = 6.18, $p < 0.0001$, $r = 0.89$. This remains a significant difference if we only take digit migrations in the 15 4-digit numbers (0.3%) and compare them to letter migrations in the 4–5 letter words (9.5%), *Wilcoxon z* = 6.11, $p < 0.0001$, $r = 0.89$ (as we have seen above, there is no length effect in migrations in word reading, and the

---

[3]We defined a target word as "clearly more frequent than its transposition counterpart" if [number of judges who judged the target as more frequent/(2*number of judges who judged the response as more frequent + number of judges who judged the target and response as having similar frequency)] was larger than 1. We followed the same procedure for selecting response words as clearly more frequent.

[4]A different level of familiarity, familiarity at the bigram level, did not seem to affect the participants' errors. We calculated the bigram frequency of each of the target words in the screening test and the migratable word list that yielded a transposition response, and the bigram frequency for the transposition responses (bigram frequencies taken from Sak et al., 2008). There was no significant difference in bigram frequencies between the target words and their migration responses, $t(452) = 0.70$, $p = 0.48$.

rate of migrations is identical in 4-, 5-, and 6 letter words). The comparison of digit migrations to letter migrations in non-words yielded a similar result: the LPD group made significantly more letter migrations in migratable non-words ($M = 18.38$, $SD = 0.10$) than digit migrations in numbers ($M = 0.38$, $SD = 0.01$), Wilcoxon $z = 6.15$, $p < 0.0001$, $r = 0.89$.

There was no significant difference between the rate of digit migrations in the LPD group ($M = 0.13$, $SD = 0.33$) and in the control group ($M = 0.22$, $SD = 0.42$) (in fact, the LPD group even had a slightly smaller digit migration rate than the controls), $t(50) = 1.07$, $p = 0.28$. As 21 of the LPD participants made no digit migrations at all and 3 LPD participants made a single digit migration, none of the LPD participants differed from the control group in number reading.

DISCUSSION

This study identified a first specific type of developmental dyslexia in Turkish, developmental LPD. This is the first report of LPD in Turkish, and it joins reports of LPD in Hebrew, Arabic, and English (Friedmann and Gvion, 2001, 2005; Friedmann and Rahamim, 2007, 2014; Friedmann et al., 2010, 2015; Friedmann and Haddad-Hanna, 2012, 2014; Kohnen et al., 2012; Kezilas et al., 2014), enriching our understanding of this dyslexia and its characteristics. It also joins a growing body of evidence showing that not only acquired dyslexia, but developmental dyslexia also has various types (for reviews see Marshall, 1984; Castles and Coltheart, 1993; Castles, 2006; Coltheart and Kohnen, 2012; Castles et al., 2014; Hanley, 2017; Friedmann and Coltheart, 2018).

Turkish Dyslexics Do Make Errors in Reading Aloud, Once the Appropriate Stimuli Are Presented

It is especially interesting that this dyslexia was found in Turkish, in light of the fact that researchers of dyslexia in Turkish claim that dyslexia only manifests itself in fluency impairments (Erden et al., 2002; Özmen, 2005). In fact, Raman (2011) and Ergül (2012) claimed that Turkish-speaking children with dyslexia read accurately (their accuracy was age-appropriate) but their reading fluency is below the normal level. These suggestions follow a tradition of dyslexia research suggesting that in languages with transparent/consistent orthographies, individuals with dyslexia do not make errors but can only be detected on the basis of their lower reading speed (e.g., Wimmer, 1993).

We believe that the generalization that dyslexic readers of transparent orthographies do not make reading errors is a misconception. First, the transparency of an orthography should only affect the rate of errors in oral reading in cases of surface dyslexia. Namely, individuals with an impairment in the lexical route, who are forced to read words via the sublexical route, are expected to make fewer errors in reading words if the grapheme-to-phoneme conversion is consistent and often provides the correct reading. However, crucially, surface dyslexia is only one of 21 types of dyslexia, and the orthographic depth of a language

is not expected to affect the other types of dyslexia. Additionally, different dyslexias yield different types of errors and are affected by different dimensions of words. Therefore, to identify each type of dyslexia, the relevant stimuli need to be presented, otherwise, the person with dyslexia will not make reading errors. Therefore, for example, to detect surface dyslexia, one needs to present irregular words; to detect phonological dyslexia, one needs to present non-words, and to detect deep dyslexia one needs to present function words, abstract words, and morphologically complex words. To detect LPD, one needs to present migratable words, i.e., words in which letter migration creates other existing words. And indeed in this study, we presented to the Turkish reading dyslexics migratable words and non-words and they made migration errors in reading, sometimes even up to 30% of the words (on migratable words that allowed for consonant migration) and to 47% of the migratable non-words.

The participants made far more errors on migratable words (and non-words) than on non-migratable words (and non-words). In fact, they did not make migrations on the non-migratable words. This finding is in line with the lexical tendency that the participants showed in their migration responses: most of their migration responses (in reading the migratable word and the migratable non-words) were lexical. This means that the diagnosis of LPD in Turkish critically hinges on the types of words that are presented to the participant: if migratable words are not presented, the participant's LPD may be missed.

Through the analyses of these migration errors, the study examined, in detail, the characteristics of LPD and the way it manifested itself in Turkish. We found that many of the characteristics of LPD reported for other languages held also for our Turkish-speaking participants more middle than exterior letter migrations, slightly more adjacent than non-adjacent migrations, and we were also able to discover new properties, which the special nature of the Turkish language and orthography allowed us to examine. Below, we report and discuss the main properties of LPD in Turkish that emerged from this study.

No Effect of the Morphological Structure of the Word

Turkish has very rich morphology, which made it a wonderful testing ground for examining the effect of morphology on letter position encoding. We examined several points regarding the interaction of morphology and letter position encoding. First, we asked whether more migrations occurred in morphologically complex words compared with morphologically simple ones. The results were that there were no differences between the rates of migrations in morphologically complex and morphologically simple words.[5] In addition, many of the migrations changed the morphological structure of the target word. The migrations were also insensitive to whether the letters belonged to the stem or

---

[5]This pattern differs from other types of dyslexia where morphologically complex words are more prone to errors. For example, Çapan (1989) reports on two Turkish children with dyslexia who made more errors on longer words and made many omissions in morphologically complex words. Their deficit may have been in the orthographic input buffer or the phonological output buffer, stages that are sensitive to the morphological structure of the target word (Sternberg and Friedmann, 2007; Dotan and Friedmann, 2015).

the affix: the participants actually made more transpositions of letters of the stem with letters from the non-stem morpheme than within-morpheme migrations.

These findings suggest, in line with Friedmann et al. (2015), that letter position encoding precedes morphological analysis, a conclusion that is quite sensible: to perform morphological analysis, the system needs to know first exactly where each letter is localized.

Migrations Are Sensitive to the Consonant-Vowel Status of the Migrating Letters

In a recent paper, Khentov-Kraus and Friedmann (2018) reported on vowel dyslexia, which selectively affects the reading of vowel letters. This dyslexia results from a selective deficit in the processing of vowel letters in the sublexical route. In the framework of that paper, the researchers also analyzed migration errors of 48 Hebrew readers with LPD and found that they make more errors that involve the transposition of two consonants than transpositions of a vowel and a consonant. This, in turn, was taken to suggest that the orthographic-visual analyzer is already sensitive to the consonant-vowel status of the target letters. In the current study we took this examination one step further, by using the opportunities offered to us by the Turkish language and orthography. We compared CC transpositions (transpositions of a consonant letter with another consonant letter), with CV transpositions (transpositions of a consonant letter with a vowel letter), and added a comparison that has not been done yet, of VV transpositions – transpositions that only involve two vowels exchanging positions. We did so by selecting three types of migratable words, each allowing different types of transposition.

The results were clear-cut: CC migrations occurred almost three times more than either VC or VV migrations. This finding applies also to non-words, where most of the migrations involved consonants only. This finding has a very important bearing on the orthographic-visual analyzer: it means that already at the stage of letter position encoding, the orthographic-visual analyzer is sensitive to the consonant-vowel status of the letter. Namely, even though *consonant* and *vowel* are phonological notions, the orthographic processing is sensitive to this distinction at the letter level, and distinguishes between consonant letters and vowel letters already at the orthographic-visual analysis stage, long before the phonological stages of reading.

This difference between migrations of consonant and vowels also says something about the nature of the LPD deficit: it is not a visual deficit, but rather a deficit in a later, orthographic stage. Had the deficit been visual, no difference between consonant and vowel letters would be predicted.

The finding that there were mainly CC migrations also accounts for the finding that more migrations occurred between syllables than within a syllable: syllable structure in Turkish is regular, and syllables take the forms CV, VC, CVC, and VCV. As a result, migration within a CV or VC syllable will always be CV migration, whereas migration across syllables can be CC migration. The tendency to make more CC migrations results in making more across-syllables migrations.

The Locus of the Deficit That Gives Rise to LPD

The Deficit Is in a Stage Shared by the Lexical and Non-lexical Routes: Non-word Reading

We tested the participants' reading, not only of existing words but also of non-words. This is important in order to examine whether the deficit indeed lies in the orthographic-visual analyzer or whether it stems from a deficit in the orthographic input lexicon. We found that the participants made migration errors not only on words but also on (migratable and non-migratable) non-words, and that their non-word reading shows the same error types (migrations) and the same characteristics as word reading. These findings indicate that the deficit that gives rise to LPD has to reside in a non-lexical stage that is shared by the lexical and sublexical routes.

The Deficit Is in an Input Reading Stage and Not in a Phonological Output Stage

Two such shared stages exist the orthographic-visual analyzer and the phonological output buffer. Which of them is responsible for LPD? If the deficit lies in the orthographic-visual analyzer, then the deficit should not only affect reading aloud but also other tasks that involve reading input, even without oral production. If the deficit is in the phonological output buffer, silent reading tasks should not involve migrations. To examine this, we tested the participants' same-different decision, lexical decision of migratable non-words, and the comprehension of migratable words that required distinguishing between the target word and its migration counterpart.

We found that all the LPD participants had letter migrations not only in oral reading but also in at least one of these silent reading tasks. These results support the localization of the deficit in the orthographic-visual analysis, pre-lexical stage.

To further explore this point, and examine the phonological output buffer stage, we also asked them to repeat the migratable words they had read with a migration error. Had the deficit originated in the phonological output stage, we would expect the participants to also make these errors when they repeated the same words. The results unequivocally showed that they were unimpaired in the phonological output stage – they repeated the migratable words correctly, and significantly better than their reading of the same words. We reached the same conclusion on the basis of a non-word repetition task in which these participants performed within the normal range, again, ruling out a deficit in their phonological output stage. Additionally, the finding that there was no significant length effect on migrations also supports the conclusion that the deficit does not reside in the phonological output buffer.

To conclude, then, the results indicate that the participants' deficit that gives rise to LPD is in the orthographic-visual analyzer, in the function responsible for letter position encoding. This function is already sensitive to the consonant-vowel status of the target letters but is not sensitive to the morphological structure of the target word.

Letter Positions Are Underspecified, Rather Than Incorrectly Encoded: Evidence From the Frequency Effect

The frequency had a significant effect on migrations. The participants made far more migrations when the target word was less frequent than the migration result than when the target word was more frequent than its migration result. This, too, has implications for diagnosis, as well as for the description of the LPD impairment. Clinically, it means that in order to detect LPD, it is better to present the less frequent migratable word than its more frequent counterpart. Theoretically, it provides insights as to the nature of the impaired process in LPD.

One can imagine two possibilities for the failure in letter position encoding: one is that letter identities are bound to incorrect letter positions, the other is that the position of some (usually middle) letters is not encoded. These two descriptions bear different predictions with respect to the effect of frequency on migrations: if it is erroneous letter position encoding, the input to the orthographic input lexicon is letters that appear in an incorrect order, and if this letter order exists in the lexicon, it doesn't matter how frequent it is, so we would not expect frequency to affect the errors. If, on the other hand, frequency does have an effect, as we see here, it means that letter positions are not encoded and then the lexicon is searched with this partial information, of letter identities without positions. In this case, the orthographic input lexicon finds the first lexical entry that matches the partial information, which will usually be the more frequent word. Thus, the frequency effect we detected suggests that our participants did not encode the position of some of the letters, rather than encoded it incorrectly.

Not a General Deficit in Sequence Perception: Normal Number Reading

Another question that is often raised with respect to LPD is whether it is dyslexia that affects only orthographic material or whether it is a more general perceptual deficit that also affects other sequences. To examine this, we tested our participants' reading of multi-digit numbers. We found that none of the participants had a deficit in reading numbers and none of them made more digit migrations than the controls.

This indicates, in line with other studies on LPD (Friedmann et al., 2010) and on other dyslexias in the orthographic-visual analyzer (see Dotan and Friedmann, 2019, for a review of dissociations between dyslexia and dysnumeria), that the orthographic-visual analyzer is orthographic-specific and does not handle digits. It further indicates that LPD is orthographic-specific.

Theoretical Implications for the Reading Model

These results bear theoretical implications for the word reading process. Firstly, the finding that letter migrations were unaffected by the morphological structure of the target word suggests an insight with respect to the relative order of letter position encoding and morphological analysis. It indicates that letter position encoding happens before the system can parse the

morphological structure of the target word. This makes sense, as morphological analysis needs to apply to strings of letters that are already bound to positions within the word.

A second theoretical implication regards the processing of consonant and vowel letters. The findings that consonant transposed with other consonants far more often than consonants with vowels, and that consonants migrated more than vowels, indicate that the consonant-vowel status of the letter is already computed early in the orthographic-visual analysis stage, before letter position encoding. This finding also suggests that the position of consonant and vowel letters is encoded separately. The finding that consonant-consonant migrations were far more frequent than consonant-vowel migrations, which was also found in Khentov-Kraus and Friedmann (2018) for Hebrew, can be accounted for by assuming that consonant letters and vowel letters are encoded in two separate layers – a consonant-letters layer and a vowel-letters layer, in which the letters are ordered by their position. If we assume that migrations occur more readily within a layer, this would account for more consonant-consonant than consonant-vowel migrations. However, the new finding from Turkish LPD, that there are also more consonant-consonant migrations than vowel-vowel migrations suggests that the position of consonant letters and of vowel letters is encoded not only separately, but differently. The view should probably not be that of two separate layers of consonants and vowels, with migrations occurring mainly layer-internally. It possibly suggests that the position of the consonants in the word is computed first, creating an ordered consonantal skeleton, and then each vowel letter is inserted into the consonant skeleton. Under such mechanism, LPD mainly affects the position encoding of the consonants in the consonantal skeleton.

Finally, as we summarize above, the selective position-encoding impairment, which affected letters but not digits, indicates that the orthographic-visual analyzer is orthographic-specific and does not handle digits (Friedmann et al., 2010; Dotan and Friedmann, 2019).

Clinical Implications

Research on Turkish often refers to fluency as the only reading aspect that is impaired in dyslexia, and possibly, as a result, dyslexia studies only report fluency measures. Some researchers conclude that Turkish readers with dyslexia do not make more errors than controls (e.g., Raman, 2011). This study showed that it is both possible and essential to also look at children's errors. To expose reading errors, it is crucial to present stimuli that will be sensitive to each type of dyslexia and will induce the relevant errors from the readers. In our case, it was migratable words that were presented and revealed that Turkish readers with dyslexia *do* make reading errors, once the appropriate stimuli are presented to them. Our study shows that, in order to diagnose LPD, the toolkit for diagnosis has to include migratable words. We were able to identify this dyslexia because we used the FRİGÜ screening test, which we created to be sensitive to the various types of dyslexia. To identify LPD, we included in the test 121 migratable words and 22 migratable non-words. These stimuli exposed the LPD of our participants.

In contrast, the non-migratable words did not yield migrations. This means that if we only used non-migratable words for testing, we would have missed the source of reading difficulty of our subjects.

Once the right stimuli are presented, it becomes possible to diagnose persons with dyslexia not only on the basis of their reading speed but also based on their error rates and the types of errors that they make. This would be a way to explain to the person with dyslexia what their problem is and to start targeting treatment at the impaired components.

And in fact, slow reading is not as detrimental to reading as are errors in reading. The parents who came with their children for the reading tests reported to us only the fact that the children were not reading correctly, and their concerns were about their children making errors, in reading aloud and also in understanding what they read. This applied more generally, not only for the parents of children who we eventually found to have LPD but also for children with surface dyslexia, attentional dyslexia, and vowel dyslexia.

A further clinical conclusion related to the properties of the migratable words selected for the diagnostic word list: in order to trigger more errors, they should include two adjacent middle consonants that may transpose and create another existing word, which is more frequent than the target one.

Thus, the clinical implications of the current study are: (A) look at errors and error types, and (B) use (less-frequent) migratable words in the word lists for diagnosing LPD, and, in general – include words that are sensitive to each dyslexia type in order to identify it.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher. The authors maintain the rights for the reading tests.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Ethical Guideline of the Anadolu University Ethical Committee with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Anadolu University Ethical Committee.

AUTHOR CONTRIBUTIONS

NF and SG conceived of the presented idea and designed the experiments. Both authors constructed the stimuli together and verified the analysis methods. NF supervised the project. SG carried out the experiments. Both authors discussed the results and wrote together the final manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Aksan, Y., Aksan, M., Mersinli, Ü, and Demirhan, U. U. (2016). *A Frequency Dictionary of Turkish*. London: Routledge.

Babayiğit, S., and Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: a longitudinal study. *Read. Writ.* 23, 539–568. doi: 10.1007/s11145-009-9173-y

Çapan, S. (1989). "A linguistic study of reading and writing disorders in Turkish, an agglutinative language," in *Reading and Writing Disorders in Different Orthographic Systems. NATO ASI Series (Series D: Behavioural and Social Sciences)*, Vol. 52, eds P. G. Aaron and R. M. Joshi (Dordrecht: Springer).

Castles, A. (2006). The dual route model and the developmental dyslexias. *Lond. Rev. Educ.* 4, 49–61. doi: 10.1080/13603110600574454

Castles, A., and Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition* 47, 149–180. doi: 10.1016/0010-0277(93)90003-E

Castles, A., Kohnen, S., Nickels, L., and Brock, J. (2014). Developmental disorders: what can be learned from cognitive neuropsychology? *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 369:20130407. doi: 10.1098/rstb.2013.0407

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Vis. Lang.* 15, 245–286.

Coltheart, M. (1985). "Cognitive neuropsychology and the study of reading," in *Attention and Performance XI*, eds M. I. Posner and O. S. M. Martin (Hillsdale, NJ: Erlbaum).

Coltheart, M. (1987). "Functional architecture of the language-processing system," in *The Cognitive Neuropsychology of Language*, eds M. Coltheart, G. Sartori, and R. Job (Hove: Erlbaum).

Coltheart, M., Curtis, B., Atkins, P., and Haller, M. (1993). Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychol. Rev.* 100, 589–608. doi: 10.1037//0033-295x.100.4.589

Coltheart, M., and Kohnen, S. (2012). "Acquired and developmental disorders of reading and spelling," in *The Handbook of the Neuropsychology of Language*, ed. M. Faust (Chichester: Wiley-Blackwell Publishers), 892–920. doi: 10.1002/9781118432501.ch43

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256. doi: 10.1037/0033-295X.108.1.204

Crawford, J. R., and Howell's, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *Clin. Neuropsychol.* 12, 482–486. doi: 10.1076/clin.12.4.482.7241

Dotan, D., and 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. doi: 10.1016/j.cortex.2014.08.014

Dotan, D., and Friedmann, N. (2019). Separate mechanisms for number reading and word reading: evidence from selective impairments. *Cortex* 114, 176–192. doi: 10.1016/j.cortex.2018.05.010

Durgunoğlu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Ann. Dyslexia* 52, 189–204. doi: 10.1007/s11881-002-0012-y

Ellis, A. W., Flude, B., and Young, A. W. (1987). Neglect dyslexia and the early visual processing of letters in words and nonwords. *Cogn. Neuropsychol.* 4, 439–464. doi: 10.1080/02643298708252047

Ellis, A. W., and Young, A. W. (1996). *Human Cognitive Neuropsychology*. Hillsdale, NJ: Erlbaum.

Erden, G., Kurdoğlu, F., and Uslu, R. (2002). İlköğretim okullarına devam eden Türk çocuklarının sınıf düzeylerine göre okuma hızı ve yazım hataları normlarının geliştirilmesi. *Türk Psikiyatri Dergisi* 13, 5–13.

Ergül, C. (2012). Okumada güçlük yaşayan öğrencilerin okuma performanslarının öğrenme güçlüğü riski açısından değerlendirilmesi. *Kuram ve Uygulamada Eğitim Bilimleri* 12, 2033–2057.

Friedmann, N., and Coltheart, M. (2018). "Types of developmental dyslexia," in *Handbook of Communication Disorders: Theoretical, empirical, and Applied Linguistic Perspectives*, eds A. Bar-On and D. Ravid (Berlin: De Gruyter Mouton), 721–751.

Friedmann, N., Dotan, D., and Rahamim, E. (2010). Is the visual analyzer orthographic-specific? Reading words and numbers in letter position dyslexia. *Cortex* 46, 982–1004. doi: 10.1016/j.cortex.2009.08.007

Friedmann, N., and Gvion, A. (2001). Letter position dyslexia. *Cogn. Neuropsychol.* 18, 673–696. doi: 10.1080/02643290143000051

Friedmann, N., and Gvion, A. (2003). *TILTAN: Battery for the Diagnosis of Dyslexias in Hebrew*. Tel Aviv: Tel Aviv University.

Friedmann, N., and Gvion, A. (2005). Letter form as a constraint for errors in neglect dyslexia and letter position dyslexia. *Behav. Neurol.* 16, 145–158. doi: 10.1155/2005/635634

Friedmann, N., Gvion, A., and Nisim, R. (2015). Insights from developmental and acquired letter position dyslexia on morphological decomposition in reading. *Front. Hum. Neurosci.* 9:143. doi: 10.3389/fnhum.2015.00143

Friedmann, N., and Haddad-Hanna, M. (2012). Letter position dyslexia in Arabic: from form to position. *Behav. Neurol.* 25, 193–203. doi: 10.3233/BEN-2012-119004

Friedmann, N., and Haddad-Hanna, M. (2014). "Types of developmental dyslexia in Arabic," in *Handbook of Arabic Literacy: Insights and Perspectives. Language and Literacy Series*, eds E. Saiegh-Haddad and M. Joshi (The Netherlands: Springer), 119–152.

Friedmann, N., and Rahamim, E. (2007). Developmental letter position dyslexia. *J. Neuropsychol.* 1, 201–236. doi: 10.1348/174866407X204227

Friedmann, N., and Rahamim, E. (2014). What can reduce letter migrations in letter position dyslexia? *J. Res. Read.* 37, 297–315. doi: 10.1111/j.1467-9817.2011.01525.x

Fritz, C. O., Morris, P. E., and Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *J. Exp. Psychol.* 141, 2–18. doi: 10.1037/a0024338

Güven, S., and Friedmann, N. (2014). *FRİGÜ Screening and Diagnostic Test Battery for Dyslexia in Turkish*. Tel Aviv: Anadolu Tel Aviv University.

Hanley, J. R. (2017). Is there just one dyslexic reader? Evidence for the existence of distinct dyslexic sub-groups. *Curr. Dev. Disord. Rep.* 4, 101–107. doi: 10.1007/s40474-017-0125-y

Humphreys, G., Evett, L. J., and Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cogn. Psychol.* 22, 517–560. doi: 10.1016/0010-0285(90)90012-S

Jackson, N., and Coltheart, M. (2001). *Routes to Reading Success and Failure*. Hove: Psychology Press.

Kabak, B., and Vogel, I. (2001). The phonological word and stress assignment in Turkish. *Phonology* 18, 315–360. doi: 10.1017/s0952675701004201

Kesikçi, H., and Amado, S. (2005). Okuma güçlüğü olan çocukların fonolojik bellek, kısa süreli bellek ve wisc-r testi puanlarına ait bir inceleme. *Türk Psikoloji Dergisi* 20, 99–113.

Kezilas, Y., Kohnen, S., McKague, M., and Castles, A. (2014). The locus of impairment in English developmental letter position dyslexia. *Front. Hum. Neurosci.* 8:356. doi: 10.3389/fnhum.2014.00356

Khentov-Kraus, L., and Friedmann, N. (2018). Vowel letter dyslexia. *Cogn. Neuropsychol.* 35, 223–270. doi: 10.1080/02643294.2018.1457517

Kohnen, S., Nickels, L., Castles, A., Friedmann, N., and McArthur, G. (2012). When 'slime' becomes 'smile': developmental letter position dyslexia in English. *Neuropsychologia* 50, 3681–3692. doi: 10.1016/j.neuropsychologia.2012.07.016

Marshall, J. C. (1984). "Toward a rational taxonomy of the acquired dyslexias," in *Dyslexia: A global Issue*, eds R. N. Malatesha and H. Whitaker (The Hague: Martinus Nijhoff), 211–232. doi: 10.1007/978-94-009-6929-2_14

Marshall, J. C., and Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *J. Psycholinguist. Res.* 2, 175–199. doi: 10.1007/BF01067101

Öney, B., and Durgunoğlu, A. Y. (1997). Beginning to read in Turkish: a phonologically transparent orthography. *Appl. Psycholinguist.* 18, 1–15. doi: 10.1017/S014271640000984X

Özata, H., Babür, N., and Haznedar, B. (2016). "Phonological awareness in reading acquisition," in *The Acquisition of Turkish in Childhood*, eds B. Haznedar and F. N. Ketrez (Amsterdam: John Benjamins), 243–271. doi: 10.1075/tilar.20.11oza

Özdemir, F., Haznedar, B., and Babür, N. (2012). The development of reading in early bilingualism: evidence from Turkish child L2 Learners of English. *Boğaziçi Üniversitesi Eğitim Dergisi* 29, 1–26.

Özmen, R. G. (2005). Öğrenme güçlüğü olan öğrencilerin okuma hızlarının metinlerde karşılaştırılması. *Eğitim ve Bilim* 30, 25–30.

Patterson, K. E. (1981). Neuropsychological approaches to the study of reading. *Br. J. Psychol.* 72, 151–174. doi: 10.1111/j.2044-8295.1981.tb02174.x

Raman, I. (1999). *Single-word Naming in a Transparent Alphabetic Orthography*. PhD dissertation, Middlesex University, London.

Raman, I. (2011). The role of age of acquisition in picture and word naming in dyslexic adults. *Br. J. Psychol.* 102, 328–339. doi: 10.1348/000712610X522572

Raman, I., and Weekes, B. (2008). "Cognitive neuropsychology of acquired language disorders in a Turkish-English bilingual case," in *Proceedings of the 10th International Conference on Cognitive Neuroscience*, Bodrum, doi: 10.3389/conf.neuro.09.2009.01.228

Raman, I., and Weekes, B. S. (2005). Acquired dyslexia in a Turkish-english speaker. *Ann. Dyslexia* 55, 79–104. doi: 10.1007/s11881-005-0005-8

Sak, H., Güngör, T., and Saraçlar, M. (2008). "Turkish language resources: morphological parser, morphological disambiguator and web corpus," in *Proceedings of the International Conference on Natural Language Processing*, (Berlin: Springer), 417–427. doi: 10.1007/978-3-540-85287-2_40

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.

Sternberg, T., and Friedmann, N. (2007). Developmental graphemic buffer dyslexia. *Lang. Brain* 6, 91–96.

Topbaş, S., Kaçar-Kütükçü, D., and Kopkalli-Yavuz, H. (2014). Performance of children on the Turkish nonword repetition test: effect of word similarity, word length, and scoring. *Clin. Linguist. Phonet.* 28, 602–616. doi: 10.3109/02699206.2014.927003

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Appl. Psycholinguist.* 14, 1–33. doi: 10.1017/s0142716400010122

First and Second Language Reading Difficulty Among Chinese–English Bilingual Children: The Prevalence and Influences From Demographic Characteristics

Yue Gao[1†], Lifen Zheng[1†], Xin Liu[2], Emily S. Nichols[3], Manli Zhang[4], Linlin Shang[1], Guosheng Ding[1], Xiangzhi Meng[5,6*] and Li Liu[1*]

[1] State Key Laboratory of Cognitive Neuroscience and Learning and IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China, [2] Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, [3] Department of Physics and Astronomy, University of Western Ontario, London, ON, Canada, [4] Maastricht Brain Imaging Center, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands, [5] School of Psychological and Cognitive Sciences, Beijing Key Laboratory of Behavioral and Mental Health, Peking University, Beijing, China, [6] PekingU-PolyU Center for Child Development and Learning, Beijing, China

Learning to read a second language (L2) can pose a great challenge for children who have already been struggling to read in their first language (L1). Moreover, it is not clear whether, to what extent, and under what circumstances L1 reading difficulty increases the risk of L2 reading difficulty. This study investigated Chinese (L1) and English (L2) reading skills in a large representative sample of 1,824 Chinese–English bilingual children in Grades 4 and 5 from both urban and rural schools in Beijing. We examined the prevalence of reading difficulty in Chinese only (poor Chinese readers, PC), English only (poor English readers, PE), and both Chinese and English (poor bilingual readers, PB) and calculated the co-occurrence, that is, the chances of becoming a poor reader in English given that the child was already a poor reader in Chinese. We then conducted a multinomial logistic regression analysis and compared the prevalence of PC, PE, and PB between children in Grade 4 versus Grade 5, in urban versus rural areas, and in boys versus girls. Results showed that compared to girls, boys demonstrated significantly higher risk of PC, PE, and PB. Meanwhile, compared to the 5th graders, the 4th graders demonstrated significantly higher risk of PC and PB. In addition, children enrolled in the urban schools were more likely to become better second language readers, thus leading to a concerning rural–urban gap in the prevalence of L2 reading difficulty. Finally, among these Chinese–English bilingual children, regardless of sex and school location, poor reading skill in Chinese significantly increased the risk of also being a poor English reader, with a considerable and stable co-occurrence of approximately 36%. In sum, this study suggests that despite striking differences between alphabetic and logographic writing systems, L1 reading difficulty still significantly increases the risk of L2 reading difficulty. This indicates the shared meta-linguistic skills in reading different writing systems and

the importance of understanding the universality and the interdependent relationship of reading between different writing systems. Furthermore, the male disadvantage (in both L1 and L2) and the urban–rural gap (in L2) found in the prevalence of reading difficulty calls for special attention to disadvantaged populations in educational practice.

Keywords: reading difficulty, Chinese–English bilinguals, sex differences, urban–rural gap, first language, second language

INTRODUCTION

Reading is a foundational and crucial cognitive skill for children to become participating and contributing members in the global society. However, for some children, despite having normal intelligence and adequate education, reading is a struggle rather than an enjoyment (Stevenson et al., 1982; Chan et al., 2007). An additional challenge is that children may have to learn to read a second language (L2) at the same time due to political, social, educational, or personal reasons (Gunderson et al., 2011), regardless of whether or not they are struggling with L1 reading. In light of these difficulties, both the prevalence of reading difficulty in L1 and L2 and how varying levels of L1 reading ability affect L2 reading success become important concerns for parents, educators, and researchers.

For second language learners of English, a substantial number of studies have consistently found associations between poor reading in L2 (English) and poor reading in L1, which thus far have mostly been alphabetic languages as in the case of Spanish (Lindsey et al., 2003), Italian (D'Angiulli et al., 2001), French (Deacon et al., 2009), Dutch (Morfidi et al., 2007), Hebrew (Geva and Siegel, 2000), and Korean (Wang et al., 2006). These results demonstrate an interdependent relationship between poor reading skills in L1 and L2. A number of hypotheses, such as the Linguistic Interdependence Hypothesis (Cummins, 1979, 1981), the Linguistic Coding Differences Hypothesis (LCDH, Sparks, 1995), and the Central Processing Hypothesis (Geva and Siegel, 2000) have all stated that deficits in L1 and L2 reading may share common cognitive bases or linguistic components (Geva and Siegel, 2000). Therefore, children struggling to read in their L1 may also face challenges when learning to read a foreign language as an interdependent result. However, other theories, like the Orthographic Depth Hypothesis (Katz and Frost, 1992) and the Psycholinguistic Grain Size (Ziegler and Goswami, 2005), have posited that poor L2 reading abilities may be due to inadequately meeting the demands of the L2. According to these theories, students learning to read an opaque language as an L2 may face problems as they lack the strategies and training in whole word recognition (Abu-Rabia et al., 2013; Chung et al., 2018). Similarly, pupils learning to read an L2 with a more transparent orthography might also struggle, as they are not familiar with the grapheme–phoneme correspondence rule.

Both sets of theories were shown to be plausible with the discovery of children who were experiencing reading difficulty in either purely English (L2), or in both Chinese (L1) and English (L2), who were learning these two vastly different writing systems in primary schools in Beijing (McBride-Chang et al., 2013) and

Hong Kong (Ho and Fong, 2005; Chung and Ho, 2010; McBride-Chang et al., 2013; Tong et al., 2015). In these studies, the existence of poor English only readers provided evidence of individual differences in reading in two writing systems, suggesting that L1 and L2 reading might demand different cognitive skills. In comparison, the cross-language transfer of certain reading-related skills suggested universal linguistic underpinnings for reading in two languages. Results showed that in Beijing, 40% of the poor Chinese (L1) readers were also poor English (L2) readers. This co-occurrence, i.e., the rate of poor Chinese (L1) readers also being poor English (L2) readers was significantly above the baseline level, suggesting that poor L1 reading increased the likelihood of L2 reading problems. However, studies (McBride-Chang et al., 2013; Tong et al., 2015) have shown that reading difficulty co-occurrence among children aged 8 (approximately second graders) and aged 10 years (approximately fifth graders) is different in Hong Kong, with the former being 32% (similar to the co-occurrence in Beijing) and the latter being 57%. This might be due to development: by age 8 years, though they would have had sufficient exposure to English (McBride-Chang et al., 2013), children have started to shift from "learning to read" to "reading to learn." In the latter stage, children need to use reading as a tool to build vocabulary and knowledge, thus posing greater challenges to those students in higher grades for both L1 and L2. On the other hand, this might also be due to the relatively small sample sizes in these studies (McBride-Chang et al., 2013: Age 8 children: $N = 147$; Tong et al., 2015: Grade 5 children: $N = 162$). Therefore, it is necessary to conduct a large-scale study to examine the prevalence of PC, PE, and PB, and more importantly, the risks for poor L2 reading among poor L1 readers in older children, who have more reading experience in both languages. Understanding how and to what extent L2 reading is affected as L1 reading ability varies is important.

With that being said, factors that put children at risk for reading difficulty, particularly the urban–rural gap, and those related to students' characteristics have not been well addressed. Several socio-demographic characteristics have been suggested to affect the prevalence of reading difficulty. Among them, sex differences in reading performance have been found and replicated in numerous studies (Shaywitz et al., 1990; Rutter et al., 2004; Stoet and Geary, 2013; Quinn, 2018) and shown to not be due to sampling and measurement procedures (Arnett et al., 2017), ascertainment bias (in which males are more likely to be referred for evaluation than females with equivalent reading problems) (Quinn and Wagner, 2015), nor unequal educational opportunities for females (OECD, 2016b). In addition to sex

influences, the effect of grade level is also evident. In a study of reading-related skills in native Chinese speaking children, researchers (Lei et al., 2011) found that one group of children, despite initial early deficits in phonological and morphological awareness, caught up with the peers and acquired adequate subsequent reading ability 3 years later. These results suggested that, as children enter higher grades and receive more training, their language reading ability gradually develops. Higher graders might also acquire more reading strategies and gain more reading experiences.

School location, which reflects the school's socio-economic status, also influences children's reading achievement (Xuan et al., 2019). China has experienced unprecedented economic growth in the past few decades, and rather than benefiting the urban and rural residents equally, the growth has widened the existing gap between urban and rural regions (Sicular et al., 2007). The situation makes the contrast in urban–rural schooling in China a very special case and worth more investigation and comparison.

However, effects of sex, school location, and grade level have been mostly reported in L1 reading and to the best of our knowledge, no previous studies have investigated the influences of these demographic characteristics on the prevalence of reading difficulty in English as L2 in native Chinese speakers. Further, the demographic influences on the prevalence of L1 Chinese reading difficulty are also sparse. Despite the observed male disadvantage in the prevalence of L1 Chinese reading difficulty in a few studies (Chan et al., 2007; Song and Wu, 2008; Zuo et al., 2010), compared to the abundant studies conducted in alphabetic languages, very little is known about what factors increase vulnerability for reading difficulty in Chinese. Finally, it is unclear whether the rate of co-occurrence, i.e., the chance of being a poor L2 reader among children who are identified as poor L1 readers, also demonstrates socio-demographic differences. Therefore, studying demographic characteristics, understanding the associations between these characteristics with poor reading, and expanding our understanding from poor reading in L1 to poor reading in L2 can capture more accurately and more fully the influences that shape children's bilingual reading.

Here, we answer these questions within the framework of Chinese–English bilingual reading. We were interested in examining Chinese (L1) and English (L2) reading abilities, with an individual's ability defined as knowing "how words are identified and related to spoken language processes" (Perfetti, 1985). English has an alphabetic writing system, and following the alphabetic principle, phonological cues in English words contribute greatly to reading. Therefore, weakness in phonological processing, such as phoneme decoding and grapheme–phoneme conversion, may be at the core of children's struggles with English (L1) reading (Bradley and Bryant, 1983; Snowling, 2001; Ramus, 2014). In comparison, Chinese characters consist of various radicals arranged two-dimensionally, and the phonemic information conveyed via radicals is relatively irregular or limited. Therefore, the cognitive correlates of Chinese reading difficulty may be multifaceted, with phonological processing, orthographic awareness, and visual analysis all heavily involved (Peng et al., 2017). Both the universal and unique characteristics of Chinese and English reading

make them a particularly effective pairing for examining the inter-relationship between poor reading in L1 and L2, and for testing theories proposed under the investigation of alphabetic languages. However, remarkably few studies have sought to identify the contribution of demographic characteristics to the prevalence of Chinese and English reading ability relations, thus the question remains as to whether demographic characteristics affect the relationship between L1 and L2 reading difficulty, and whether children's different backgrounds influence the prevalence of L1 and L2 reading difficulty.

The overarching goal of this study is to examine the relationship of reading difficulty in L1 and L2 in a large sample of Chinese (L1)–English (L2) bilingual children. First, we provide basic prevalence data of reading difficulty in L1 only, in L2 only, and in both L1 and L2. Second, we build a multinomial logistic regression model to compare different types of struggling readers to normal readers, and examine how grade level, sex, and school setting affects their L1 and L2 reading abilities. Finally, we address how and to what extent L1 reading difficulties significantly increase the prevalence of poor L2 reading.

MATERIALS AND METHODS

Participants

Participants' demographics are described in **Table 1**. A total of 1,824 Chinese–English bilingual students from primary schools in Beijing were assessed. For each student, we collected data on sets of variables that included demographic characteristics (age, sex, school location, grades), intelligence tests, and reading-related tests (both a Chinese reading test and an English spelling test). Valid data here refer to a dataset with complete reading-related test data and no more than 1 variable data missing from their demographic information. 1,786 of the participants provided valid data (97.92%). Among the children with valid

TABLE 1 | Demographic characteristics and descriptive statistics for all children.

	Grade 4	Grade 5
Number of children	880	906
Number of boys	460	516
Number of girls	415	390
Number of urban students	576	642
Number of rural students	304	259
Mean age in months (SD)	118.44 ± 6.11	131.00 ± 9.46
Mean non-verbal IQ in percentile (SD)	70.40 ± 22.84	67.24 ± 24.78
Mean Chinese reading score*	2,486 ± 484	2,857 ± 356
Mean English reading score[#]	10.75 ± 8.34	16.08 ± 8.94
Number and prevalence of PC	(101)11.47%	(45)5.00%
Number and prevalence of PE	(123)13.97%	(131)14.46%
Number and prevalence of PB	(48)5.45%	(34)3.75%

*PC, poor Chinese reader; PE, poor English reader; PB, poor reader in both languages. *Chinese reading score is measured by the Chinese character recognition test (in spelling format), and the score represents how many characters a child can use to make and write down a word. The full score is 3500. [#]English reading score is measured by an English word spelling test, and the score represents how many words a child can spell. The full score is 40.*

data, there were 976 boys and 805 girls, with 5 of them lacking sex information. There were 880 4th graders and 906 5th graders; 563 students from rural schools and 1,218 from urban schools, with five children lacking school location information. Among all eight schools included in this study, five are located in the downtown area [Haidian District and Chaoyang District, with GDP per capita of $24,590 and $21,442 (USD), and total GDP ranked 1st and 2nd among all 16 districts of Beijing in 2017], another three are located in rural areas [Changping district and Miyun District, with GDP per capita of $5,884 and $8,816 (USD), and total GDP ranked 8th and 13th among all 16 districts of Beijing in 2017]. Additionally, Haidian and Chaoyang Districts are both equipped with more than two public libraries, whereas there are none in Changping and Miyun Districts (Beijing Social Development Database[1]). All children started to receive formal Chinese and English instruction in Grade 1, at approximately 6 years old. Written informed consent was obtained from each participant and their parents. The institutional review board at Beijing Normal University approved the informed consent procedures.

Based on the information we gathered, in urban primary schools students receive four English classes and four Chinese classes in a week (from Monday to Friday, each class takes 45 min), and this is true for both 4th and 5th graders. Urban school students also have the opportunity to attend English-related activities outside the classroom. Similarly, rural schools also provide four English classes and four Chinese classes for the 4th and 5th graders every week. The course arrangements, in both content and frequency, are relatively similar between rural and urban areas. This is because education in China is state-run, and the Ministry of Education of People's Republic of China is the agency of the State Council that oversees education throughout the country. They "lay down requirements and create basic documents for teaching and curriculum in elementary education; organize the examination and approval of unified course materials for elementary education; and to develop high-quality education in a comprehensive manner (OECD, 2016a)." With the foundation set, policies and strategies designed by the Ministry of Education are implemented by local departments of education under its direct management. For example, to meet the basic requirements for setting up English courses in the primary schools, schools follow the principle of short courses at high frequency and ensure at least three teaching activities per week. Finally, the participants come from Han ethnicity families and schools, indicating that their home language, school language, and social language are all Chinese.

Behavioral Measures

Raven's standard progressive matrices: This test is used to assess children's non-verbal IQ by measuring their general non-verbal reasoning ability. Scoring procedures were based on the Chinese normative data (Zhang and Wang, 1985).

Chinese character recognition test (in writing format): This test (Wang and Tao, 1996) consists of 10 groups of Chinese characters at increasing levels of difficulty. Participants were

given 40 min to write down a compound word using each provided character. As there are numerous homophones in Chinese, compound word spelling compared to character reading aloud can better assess whether a child can access character meaning. Additionally, this paper–pencil test can be administered in group and was therefore suitable for large-scale assessment. This standardized test has been widely used to screen Chinese children with reading difficulty (Liu et al., 2012, 2013; McBride-Chang et al., 2013; Peng et al., 2017). The test–retest reliability of this test is 0.970 for Grade 4 and 0.984 for Grade 5.

English word spelling test: We used an English word spelling test to screen poor English readers in Chinese–English bilingual children, as has been previously done (You et al., 2011; Li et al., 2018). Forty English words chosen from primary school English textbooks for Chinese children were included, with half high-frequency words and half low-frequency words. Each word was read aloud twice and the participants were asked to write down the word on the answer sheet. This test can be administered in a group and is suitable for large-sample studies. The test–retest reliability of this test is 0.96. Moreover, to identify the capacity of the spelling test instrument, we used the Word Identification test (Woodcock, 1987) as the screening criteria in a subsample of 94 students and identified students with normal English reading ability as well as those with deficient English reading ability. We then compared the classifying results based on word identification test and classifying results based on the English spelling test, and calculated the sensitivity (93%) and the specificity (83.7%) of the English spelling test. The sensitivity and specificity were calculated based on signal detection theory [sensitivity = True Positive/(True Positive + False Negative), specificity = True Negative/(True Negative + False Positive)] (Green and Swets, 1966). This cross validation suggested that the English word spelling test is a sensitive and valid test to screen poor English readers in Chinese–English bilingual children.

Criteria for Screening Poor Readers

With parental consent obtained, children's Chinese and English reading ability was assessed by trained psychology majors in the children's classroom. The time required for the Chinese reading test varied from 20 to 40 min. The English spelling test took up to 10 min. Both tests were administered with Chinese instructions to ensure that children fully understood the requirements. All administrators passed the College English Test-Band 6 (CET-6), which is a national English standardized test evaluating the English proficiency of undergraduates and postgraduates in People's Republic of China. The order of the two tests was randomized. At the end of the testing, any questions that participants had were answered. A fairly liberal, but not unusual, criteria was used for screening poor readers (e.g., Francis et al., 2005; McBride-Chang et al., 2013; Tong et al., 2015): first, the percentile in the Raven's test score had to be above the 40th to ensure a normal IQ; second, for poor Chinese readers (PC, $N = 146$), their performance on the Chinese character recognition test had to fall at or below 25th percentile, whereas their English spelling score had to be above this level. Similarly, to define poor English readers (PE, $N = 254$), their performance in the English spelling test needed to be at or below 25th percentile while their

Chinese test score needed to be above this level. Finally, poor bilingual readers (PB, $N = 82$) refer to those whose English and Chinese scores were both falling at or below 25th percentile. Normal readers refer to the students whose Chinese and English reading performance were both above the 25th percentile (NR, $N = 1,299$).

Data Analysis

First, we computed the prevalence of different types of poor reading (PC, PE, and PB) among Chinese–English bilingual children, for boys versus girls, for urban versus rural, and for Grade 4 versus Grade 5. Second, we applied a multinomial logistic regression, a model that allows for more than two categories of the outcome variables to be predicted, to investigate the association between reading difficulty in L1 and L2 and demographic characteristics. Reading group membership (PC versus NR, PE versus NR, PB versus NR, PC versus PE, PC versus PB, PB versus PE) was used as the criterion variable with demographic characteristics (sex, school location, grades) entered as predictors. Third, we referred to the formula developed by McBride-Chang et al. (2013) to describe the chances of a poor L1 reader concurrently manifesting difficulty in L2 reading (here referred to as co-occurrence): number of poor bilingual readers/(number of poor bilingual readers + number of poor Chinese readers) * 100%. The percentage represents the portion of readers who manifested reading difficulty in L1 and L2, among the population with L1 reading difficulty. For the baseline level of L2 reading difficulty, we referred to a second formula (McBride-Chang et al., 2013): number of poor English readers/(number of poor English readers + number of normal readers). The baseline rate represents the percentage of poor L2 only readers among the students without L1 reading difficulty. Based on this, the 2 * 2 contingency table was set to compare the frequencies and the co-occurrence was compared with the baseline level of L2 reading difficulty via a non-parametric test (χ^2) to examine whether L1 reading difficulty would generate a significantly higher occurrence of L2 reading difficulty in boys or girls, in urban or rural areas, and in Grade 4 or Grade 5, respectively.

RESULTS

Table 1 and **Figures 1**, **2** present the basic prevalence data for different types of reading difficulty in this sample. The prevalence of PE and PB kept stable in Grades 4 and 5. In contrast, there was a drop in the prevalence of PC in Grade 5 compared to Grade 4. To obtain an overall picture of reading difficulty in Beijing, we collapsed data from the fourth and fifth graders. Results showed that the PB prevalence (4.60%) is lower than the PE (14.22%) and PC (8.20%) prevalence across the whole sample.

Next, we examined the gender, grade, and school location effect on the prevalence of different types of reading difficulty (**Table 2**). The -2 log likelihood (138.362) and Chi-squared statistics ($\chi^2 = 120.0.53$, $p < 0.001$) showed that these three predictor variables provided a significant fit to the model. These variables significantly distinguished between the PC and NR groups, between the PE and NR groups, and between the PB and

NR groups. In differentiating the PC group from NR group, a one-unit increase in the sex (boys) increased the odds of being in the PC group rather than the NR group by $1/0.521 = 1.919$; a one-unit increase in grade (Grade 4) increased the odds of being in the PC group rather than the NR group by $1/0.416 = 2.404$. In differentiating the PE group from the NR group, a one-unit increase in sex (boys) increased the odds of being in the PE group rather than the NR group by $1/0.444 = 2.252$; a one-unit increase in school location (rural) increased the odds of being in the PE group rather than the NR group by 2.173. In differentiating the PB group from the NR group, a one-unit increase in sex (boys) increased the odds of being in the PB group rather than the NR group by $1/0.442 = 2.262$; a one-unit increase in the grade (Grade 4) increased the odds of being in the PB group rather than the NR group by $1/0.588 = 1.701$. In differentiating among PC, PB, and PE groups, results consistently showed that the 4th graders were at higher risk than the 5th graders (2.293 times higher in PE than PC, 1.696 times higher in PE than PB), and students in the rural areas were at higher risk than their urban counterparts (1.748 times higher in PE than PC, 1.624 times higher in PE than PB) in manifesting L2 reading difficulty. In summary, being a boy significantly increased the odds of falling into all the three reading difficulty groups. Compared to being in the 5th grade, being in the 4th grade significantly increased the chances of being identified as a poor reader. Being from a rural community particularly influenced poor English reading.

To examine whether and to what extent L1 reading difficulties significantly increase the prevalence of performing poorly in L2 reading, we computed the co-occurrence (the chance of poor Chinese readers also being poor English readers) and the baseline level of L2 reading difficulty in the control group (**Table 3**). No significant sex, location, or grade differences were observed in co-occurrence levels. Co-occurrence was significantly higher than the baseline level regardless of sex ($\chi^2 = 13.78$, $p < 0.001$ in boys; $\chi^2 = 45.44$, $p < 0.001$ in girls), school locations ($\chi^2 = 50.70$, $p < 0.001$ in urban schools; $\chi^2 = 3.29$, $p = 0.081$ a marginally significant effect in rural schools), and grades ($\chi^2 = 18.34$, $p < 0.001$ in Grade 4; $\chi^2 = 35.81$, $p < 0.001$ in Grade 5).

DISCUSSION

In a large epidemiological sample of Beijing primary school children in Grade 4 and Grade 5, we investigated the prevalence of reading difficulty in L1 (Chinese) and L2 (English) children, and the influence of grade, sex, and school location on reading difficulty. There were three major findings: (1) The co-occurrence rate was significantly higher than baseline levels regardless of sex, school location, and grades. This indicates that being a poor reader in Chinese (L1) significantly increases the risk of also becoming a poor English reader. (2) In general, girls were better at reading in both Chinese and English, shown by the lower risks of all three types of poor reading (PC, PE, and PB). (3) A rural and lower grade disadvantage was observed particularly in PE.

Our first finding, that being a poor reader in Chinese significantly increases the risk of also being a poor English reader, supports theories arguing that deficits in L1 and L2 reading
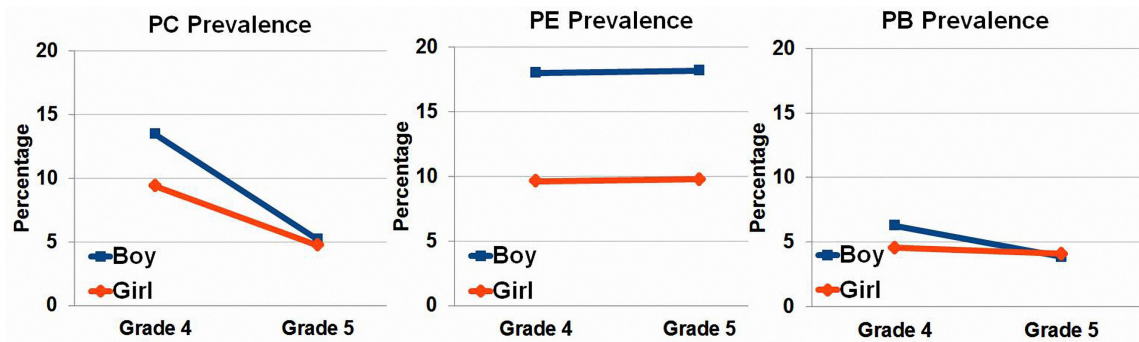
FIGURE 1 | PC, PE, and PB prevalence of boys versus girls, illustrated in Grade 4 and Grade 5 separately. PC, poor Chinese reading; PE, poor English reading; PB, poor bilingual reading.
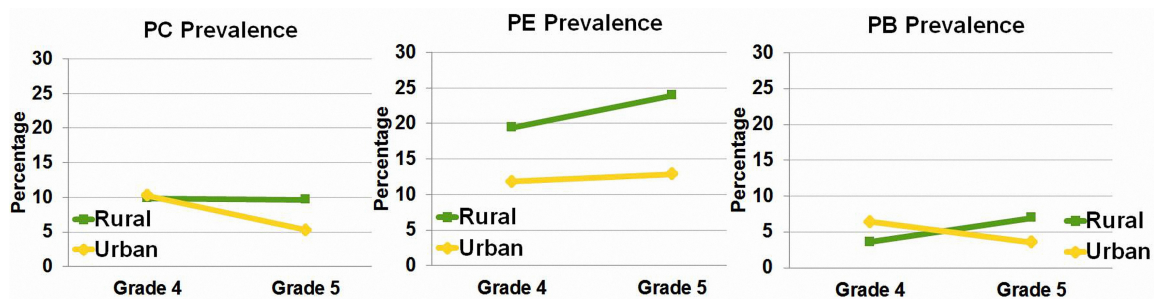

FIGURE 2 | PC, PE, and PB prevalence of rural schools versus urban schools, illustrated in Grade 4 and Grade 5 separately. PC, poor Chinese reading; PE, poor English reading; PB, poor bilingual reading.

might share some common bases (Cummins, 1979, 1981) or linguistic components (Geva and Siegel, 2000), and involves cross-linguistic transfer from L1 to L2 (Chung and Ho, 2010). We observed a stable co-occurrence of L1 and L2 reading difficulty in Chinese–English bilingual children. Across the sexes and different school locations, the chances of a poor L1 reader showing L2 reading difficulty at the same time were approximately 36%. This probability is not influenced by sex (boys: 35.51%, girls: 36.67%), school location (urban: 36.81%; rural: 34.52%), or grade (Grade 4, 32.21%; Grade 5: 42.50%). The co-occurrences observed are similar to the 40% probability reported in a relatively small sample ($N = 291$, age 8 years) of Beijing primary school children (McBride-Chang et al., 2013), as well as the 32% co-occurrence rate reported in Hong Kong primary school children ($N = 147$, age 8 years) (McBride-Chang et al., 2013). Combining these results with our current study, the co-occurrence of being poor readers in L1 and in L2 appears to be relatively stable in primary school children who learn to read English in Beijing. However, the co-occurrence we observed is smaller than the 57% co-occurrence rate reported among 5th graders in Hong Kong (Tong et al., 2015). Regarding this 57% rate, researchers reported that the sample came from several schools in which children may have been taught similar learning skills or were exposed to similar teaching methods, perhaps increasing the overlap in reading difficulties for Chinese and English (Tong et al., 2015). Alternatively, the observed difference

between the co-occurrence of Beijing students and Hong Kong students may reflect the fact that poor reader status in Chinese and English across Beijing and Hong Kong is somehow different (McBride-Chang et al., 2013).

The co-occurrence we observed is also different from what has been found among English–Spanish bilingual children (55%) (Manis and Lindsey, 2010). This difference may be attributed to a higher similarity between English and Spanish. These two languages both have alphabetic writing systems, enabling deficits in one language to be easily transferred to the other. In contrast, as Chinese is a logographic language, cross-linguistic transfer between Chinese and English might be relatively weaker compared to that between Spanish and English. The hypothesis of low cross-linguistic transfer is supported by the small correlations found between Chinese and English reading-related cognitive skills (Yang et al., 2017). Additionally, another study (Pasquarella et al., 2015) examining the cross-language transfer of word reading in Spanish–English and Chinese–English bilinguals found that transfer of word reading accuracy is based on the structural similarities between the L1 and L2 scripts.

On the one hand, our finding of a 36% co-occurrence in reading difficulty across Chinese and English suggests that in addition to the assessment of L1 reading skill, the assessment of L2 reading skills is critical in early L2 readers, as a poor reader in one language may not necessarily be a poor reader in another language. On the other hand, this finding suggests that we need

TABLE 2 | Multinomial logistic regression.

	NR versus PB			NR versus PC			NR versus PE			PC versus PE			PC versus PB			PB versus PE		
	B(SE)	Exp(B)	95% CI	B(SE)	Exp(B)	95% CI	B(SE)	Exp(B)	95% CI	B(SE)	Exp(B)	95% CI	B(SE)	Exp(B)	95% CI	B(SE)	Exp(B)	95% CI
Grade	**−0.531**** **(0.199)**	0.588	0.398 0.868	**−0.876***** **(0.169)**	0.416	0.299 0.580	−0.003 (0.132)	0.997	0.770 1.291	**0.873***** **(0.195)**	2.393	1.633 3.508	0.345 (0.246)	1.411	0.872 2.284	**0.528*** **(0.221)**	1.696	1.099 2.617
Sex	**−0.815***** **(0.206)**	0.442	0.296 0.662	**−0.653***** **(0.165)**	0.521	0.377 0.719	**−0.812***** **(0.136)**	0.444	0.340 0.580	−0.159 (0.196)	0.853	0.581 1.252	−0.163 (0.248)	0.850	0.522 1.383	0.004 (0.231)	0.987	0.638 1.579
School	0.291 (0.209)	1.338	0.888 2.015	0.218 (0.173)	1.243	0.886 1.745	**0.776***** **(0.135)**	2.173	1.668 2.832	**0.559**** **(0.196)**	1.748	1.190 2.569	0.073 (0.252)	1.076	0.657 1.763	**0.485*** **(0.228)**	1.624	1.038 2.542

95% CI, 95% confidence intervals for Exp B; PC, poor Chinese reader; PE, poor English reader; PB, poor reader in both languages; SE, standard error; Exp(B) refers to the odds ratios. $p < 0.05$, $p < 0.01$, $p < 0.001$. The bolded values refer to significant effects.

to pay additional close attention to L2 reading ability of poor Chinese (L1) readers, as reading difficulty in Chinese increased the possibility of being poor readers in English (L2).

Our second finding of sex differences in co-occurrence of reading difficulty is consistent with previous studies. The sex imbalance in dyslexia prevalence is well documented in alphabetic languages with a sex ratio ranging from approximately 3:1 to 5:1 in referred samples and from 1.5:1 to 3.3:1 in epidemiological samples (Shaywitz et al., 1990; Rutter et al., 2004). Recently, a meta-analysis including 16 studies ($N = 552,729$) concluded that males are more likely than females to be identified as having reading difficulties regardless of methodological and statistical influences (Quinn, 2018). Similarly, a within- and across-nation assessment of 10 years of Programme for International Student Assessment (PISA) data (Stoet and Geary, 2013) confirmed a male disadvantage in reading. The increased prevalence of dyslexia in boys versus girls has also been reported in China with a sex ratio from 1.6:1 to 2.0:1 in Cantonese speaking children in Hong Kong (Chan et al., 2007), and from 1.8:1 to 2.45:1 in Mandarin speaking children in Mainland China (Song and Wu, 2008; Zuo et al., 2010) in epidemiological samples. The sex gap could be due to sex differences in cognition (Kimura, 1999; Halpern, 2000), learning strategy (Poole, 2005; Griva et al., 2012), attitude toward second language learning (Davies, 2004), or a complex gene–environment interaction (Van Der Slik et al., 2015).

Our epidemiological data extend these previous studies by showing that the higher prevalence of reading difficulty in boys compared to girls was not only in L1, but also in L2. In fact, the sex difference was even more pronounced in L2 than in L1. Research on the impact of sex on L2 acquisition is much scarcer than on sex effects in L1 acquisition. Burstall (1975) reported that girls scored significantly higher than boys in learning French as a second language from age 13 to age 16 years. Davies (2004) further showed that this sex gap actually started as early as age 7 years, the first term when children started to learn French. The sex gap

TABLE 3 | The co-occurrence of L1 and L2 poor reading and the baseline rate of L2 poor reading in different school locations, sex, and grades.

	Co-occurrence	L2 poor reading	χ^2
Boys	35.51%	26.78%	13.78***
Girls	36.67%	12.07%	45.44***
Ratio	0.97:1	2.22:1	
χ^2	0.03		
Urban	36.81%	16.54%	50.70***
Rural	34.52%	33.52%	3.29
Ratio	1.07:1	0.49:1	
χ^2	0.13		
Grade 4	32.21%	20.40%	18.34***
Grade 5	42.50%	18.82%	35.81***
Ratio	0.76:1	1.08:1	
χ^2	2.63		

*Co-occurrence refers to the chances of becoming a poor reader in English given that the child was already a poor reader in Chinese. It was calculated by the formula (McBride-Chang et al., 2013): N of PB/(N of PB + N of PC) * 100%. The baseline rate of L2 poor reading was calculated by the formula (McBride-Chang et al., 2013): N of PE/(N of PE + N of NR). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.*

has also been reported in Korean learning English as a foreign language (Pae, 2004) and in adult learners of Dutch as a second language across countries of origin and continents (Van Der Slik et al., 2015). In terms of Chinese learners of English, Boyle (1987) reported that female Chinese students outperformed their male counterparts in English listening skills. To our knowledge, the current study is the first to report a significantly higher ratio of poor L2 reading in boys versus girls in Chinese children learning English as a second language.

Our third finding is the rural disadvantage in the prevalence of reading difficulty in L2 (English), with a larger disparity in L2 than in L1 reading. The urban–rural gap in reading performance has been frequently observed in large-scale studies related to L1 across countries such as PISA and Progress in International Reading Literacy Study (PIRLS) (Cartwright and Allen, 2002). These studies did not yield a consistent rural disadvantage, but revealed that rural–urban gap in reading performance in L1 varied in direction and magnitude across countries (Cartwright and Allen, 2002; Elijio and Urban, 2006). Most importantly, studies (Young, 1994; Cartwright and Allen, 2002; Elijio and Urban, 2006) have suggested that several factors contribute to the link between school location and reading performance. These factors include socio-economic environment (Elijio and Urban, 2006), school educational quality (Elijio and Urban, 2006), community differences in levels of adult education (Cartwright and Allen, 2002), and school–community connection (Tharp and Gallimore, 1991). One study (Wang et al., 2018) found that for Chinese primary students, the observed rural–urban reading literacy gap in L1 is mediated by parental education level and family literacy environment. In our study, the larger rural–urban gap in L2 compared to L1 reading abilities may reflect the fact that L2 reading is more susceptible to these above-mentioned factors. Moreover, children in lower grades are more vulnerable to L2 reading difficulty when compared to L1 reading difficulty, indicating that children's L2 reading skill might be more influenced by the environment. One study (Kieffer, 2011) found that English L2 learners in the United States with initially limited English proficiency demonstrated English reading trajectories that were below national averages, but converge with peers from similar socioeconomic backgrounds after elementary school. Therefore, for children learning English as second language in China, those attending urban schools are more likely to have access to and benefit from more abundant L2 learning resources in family and school. Additionally, as children enter higher grades, they might receive more targeted tutoring in English education programs than do those who are in lower grades. In sum, the rural and lower grade disadvantage in the prevalence of reading difficulty in L2 observed in the current study could be attributed to one or more of these above factors, but will require more investigation to further our understanding.

Amplified by the Matthew Effect, the concept arising from findings that individuals who have advantageous early educational experiences are able to utilize new educational experiences more efficiently (Walberg and Tsai, 1983; Stanovich, 1986), the widening rural–urban gap is concerning. Poor readers in urban areas are more likely to be noticed, assessed, and to receive intervention with the help from well-educated parents,

qualified teachers, and well-resourced urban settings. Studies have shown the significant education inequality of urban–rural area in China (Zhang et al., 2015), and rural children might be less likely to receive targeted instruction from the rural educational systems. These factors may lead to imbalances in the developmental trajectory of reading abilities between urban and rural areas, thus enlarging the disparity.

Limitations and Future Directions

A number of caveats need to be noted regarding the present study. First, we used a lower end cutoff score of 25% to define reading difficulty; however, this arbitrary cutoff score approach has been critiqued for lack of stability over time (Francis et al., 2005). Nevertheless, the cutoff score approach remains one of the most common ways to define reading difficulties (e.g., Manis and Lindsey, 2010). Moreover, previous studies (McBride-Chang et al., 2013; Tong et al., 2015) investigating the prevalence of poor English reading in native Chinese speaking children also adopted this approach. Second, we aimed to recruit a representative sample of Beijing primary school children. For example, based on recommendations of local education administration officers, we recruited children from urban and rural schools with different levels of teaching quality and educational environment. Despite these efforts, the representativeness of our sample is open to discussion as we did not apply a completely random sampling approach. Finally, we admit that an English reading test would be the ideal instrument in screening English poor readers, but we only implemented a word spelling test. The reasons are that in China, English is mostly learned as a second language and we cannot directly use the standardized English reading tests developed for native English populations. Considering the practical issues and the time constraints imposed by the participating elementary schools, we used an English spelling test, which can be administered in large-scale settings and also can provide reliable and valid data regarding children's current English reading abilities. Nevertheless, a richer set of reading tests as well as reading-related cognitive skill tests are needed in future research. Other psycholinguistic factors, for example, the age of acquisition should also be investigated to better depict the reading and cognitive profiles of reading difficulties in different writing systems such as Chinese and English (Davies et al., 2017). Employing multiple measurements as a means of identifying learners and assessing progress or future needs is recommended in order to develop a complete profile of a bilingual's L1 and L2 language reading challenges.

CONCLUSION

The overarching conclusion of the present study is that in Chinese–English bilingual children, despite striking differences between alphabetic and logographic writing systems, L1 reading difficulty still significantly increases the risk of L2 reading difficulty. This supports theories arguing for shared linguistic components in reading different writing systems, and underlines the importance of understanding the universality of reading

between different writing systems. Furthermore, the male disadvantage and the urban–rural gap in the prevalence of reading difficulty call for special attention from the educational system and policy makers. These conclusions are only preliminary, and the need for more rigorous research of disadvantaged groups is evident.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Beijing Normal University Ethics Committee.

Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

REFERENCES

Abu-Rabia, S., Shakkour, W., and Siegel, L. (2013). Cognitive retroactive transfer (crt) of language skills among bilingual Arabic-English readers. *Biling. Res. J.* 36, 61–81. doi: 10.1080/15235882.2013.775975

Arnett, A. B., Pennington, B. F., Peterson, R. L., Willcutt, E. G., DeFries, J. C., and Olson, R. K. (2017). Explaining the sex difference in dyslexia. *J. Child Psychol. Psychiatry* 58, 719–727. doi: 10.1111/jcpp.12691

Boyle, J. P. (1987). Sex differences in listening vocabulary. *Lang. Learn.* 37, 273–284. doi: 10.1111/j.1467-1770.1987.tb00568.x

Bradley, L., and Bryant, P. E. (1983). Categorizing sounds and learning to read – a causal connection. *Nature* 303, 419–421. doi: 10.1038/301419a0

Burstall, C. (1975). Primary French in the balance. *Educ. Res.* 17, 193–198. doi: 10.1080/0013188750170304

Cartwright, F., and Allen, M. K. (2002). *Understanding the Rural-Urban Reading Gap. Education, Skills and Learning Research Papers*. Ottawa: Statistics Canada.

Chan, D. W., Ho, C. S. H., Tsang, S. M., Lee, S. H., and Chung, K. K. (2007). Prevalence, gender ratio and gender differences in reading-related cognitive abilities among Chinese children with dyslexia in Hong Kong. *Educ. Stud.* 33, 249–265. doi: 10.1080/03055690601068535

Chung, C., Chen, X., and Geva, E. (2018). Deconstructing and reconstructing cross-language transfer in bilingual reading development: an interactive framework. *J. Neurolinguist.* 50, 149–161. doi: 10.1016/j.jneuroling.2018.01.003

Chung, K., and Ho, C. S. H. (2010). Second language learning difficulties in Chinese children with dyslexia: what are the reading-related cognitive skills that contribute to English and Chinese word reading? *J. Learn. Disabil.* 43, 195–205. doi: 10.1177/0022219409345018

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Rev. Educ. Res.* 49, 222–251. doi: 10.3102/00346543049002222

Cummins, J. (ed.) (1981). "The role of primary language development in promoting educational success for language minority students," in *California Office of Bilingual Bicultural Education, Schooling and language minority students: A theoretical framework*, (Sacramento: California Department of Education), 3–49.

D'Angiulli, A., Siegel, L. S., and Serra, E. (2001). The development of reading in English and Italian in bilingual children. *Appl. Psycholinguist.* 22, 479–507. doi: 10.1017/s0142716401004015

Davies, B. (2004). The gender gap in modern languages: a comparison of attitude and performance in year 7 and year 10. *Lang. Learn. J.* 29, 53–58. doi: 10.1080/09571730485200111

Davies, R. A. I., Arnell, R., Birchenough, J. M. H., Grimmond, D., and Houlson, S. (2017). Reading through the life span: individual differences in psycholinguistic effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 43, 1298–1338. doi: 10.1037/xlm0000366

Deacon, S. H., Wadewoolley, L., and Kirby, J. R. (2009). Flexibility in young second-language learners: examining the language specificity of orthographic processing. *J. Res. Read.* 32, 215–229. doi: 10.1111/j.1467-9817.2009.01392.x

Elijio, A., and Urban, R. A. I. (2006). "Reading achievements in urban and rural communities: a comparative analysis of equity in education," in *Proceedings of the Second IEA International Research Conference*, Washington, D.C.

Francis, D. J., Fletcher, J. M., Catts, H. W., and Tomblin, J. B. (2005). "Dimensions affecting the assessment of reading comprehension," in *Children's Reading Comprehension and Assessment*, eds S. G. Paris, and S. A. Stahl, (Mahwah, NJ: Lawrence Erlbaum Associates, Inc.), 369–394.

Geva, E., and Siegel, L. S. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Read. Writ.* 12, 1–30.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*, Vol. 1. New York, NY: Wiley.

Griva, E., Alevriadou, A., and Semoglou, K. (2012). Reading preferences and strategies employed by primary school students: gender, socio-cognitive and citizenship issues. *Int. Educ. Stud.* 5, 24–34.

Gunderson, L., D' Silva, R., and Chen, L. (2011). "Second language reading disability: international themes," in *Handbook of Reading Disability Research*, eds A. McGill-Franzen, and R. L. Allington, (New York, NY: Routledge), 13–24.

Halpern, D. F. (2000). *Sex Differences in Cognitive Abilities*, 3rd Edn. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Ho, C. S. H., and Fong, K. M. (2005). Do Chinese dyslexic children have difficulties learning English as a second language? *J. Psycholinguist. Res.* 34, 603–618. doi: 10.1007/s10936-005-9166-1

Katz, L., and Frost, R. (1992). "Reading in different orthographies: the orthographic depth hypothesis," in *Orthography, Phonology, Morphology, and Meaning*, eds R. Frost, and L. Katz, (Amsterdam: Elsevier), 67–84. doi: 10.1016/s0166-4115(08)62789-2

Kieffer, M. J. (2011). Converging trajectories: reading growth in language minority learners and their classmates, kindergarten to grade 8. *Am. Educ. Res. J.* 48, 1187–1225. doi: 10.3102/0002831211419490

Kimura, D. (1999). *Sex and Cognition*. Cambridge, MA: MIT press.

Lei, L., Pan, J., Liu, H., McBride-Chang, C., Li, H., Zhang, Y., et al. (2011). Developmental trajectories of reading development and impairment from ages 3 to 8 years in Chinese children. *J. Child Psychol. Psychiatry* 52, 212–220. doi: 10.1111/j.1469-7610.2010.02311.x

Li, H. H., Booth, J. R., Bélanger, N. N., Feng, X. X., Tian, M. Y., Xie, W. Y., et al. (2018). Structural correlates of literacy difficulties in the second language: evidence from mandarin-speaking children learning English. *NeuroImage* 179, 288–297. doi: 10.1016/j.neuroimage.2018.06.037

Lindsey, K. A., Manis, F. R., and Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *J. Educ. Psychol.* 95, 482–494. doi: 10.1037/0022-0663.95.3.482

Liu, L., Wang, W., You, W., Li, Y., Awati, N., Zhao, X., et al. (2012). Similar alterations in brain function for phonological and semantic processing to visual characters in chinese dyslexia. *Neuropsychologia*, 50, 2224–2232. doi: 10.1016/j.neuropsychologia.2012.05.026

Liu, L., You, W., Wang, W., Guo, X., Peng, D., and Booth, J. (2013). Altered brain structure in chinese dyslexic children. *Neuropsychologia* 51, 1169–1176. doi: 10.1016/j.neuropsychologia.2013.03.010

Manis, F. R., and Lindsey, K. A. (2010). *Cognitive and Oral Language Contributors to Reading Disabilities in Spanish-English bilingual Children. Language and Literacy Development in Bilingual Set.* New York, NY: Guilford, 280–303.

McBride-Chang, C., Shu, H., Chan, W., Wong, T., Wong, M. Y., Zhang, Y., et al. (2013). Poor readers of Chinese and English: overlap, stability, and longitudinal correlates. *Sci. Stud. Read.* 17, 57–70. doi: 10.1080/10888438.2012.689787

Morfidi, E., Leij, A. V. D., Jong, P. F. D., Scheltinga, F., and Bekebrede, J. (2007). Reading in two orthographies: a cross-linguistic study of Dutch average and poor readers who learn English as a second language. *Read. Writ.* 20, 753–784. doi: 10.1007/s11145-006-9035-9

OECD (2016a). *Education in China: A Snapshot*. Paris: OECD Publishing.

OECD (2016b). *PISA 2015 Results (Volume I): Excellence and Equity in Education.* Paris: OECD Publishing.

Pae, T. I. (2004). Gender effect on reading comprehension with Korean EFL learners. *System* 32, 265–281. doi: 10.1016/j.system.2003.09.009

Pasquarella, A., Chen, X., Gottardo, A., and Geva, E. (2015). Cross-language transfer of word reading accuracy and word reading fluency in spanish-english and chinese-english bilinguals: script-universal and script-specific processes. *J. Educ. Psychol.* 107, 96–110. doi: 10.1037/a0036966

Peng, P., Wang, C., Tao, S., and Sun, C. (2017). The deficit profiles of chinese childrenwith reading difficulties: a meta-analysis. *Educ. Psychol. Rev.* 29, 513–564. doi: 10.1007/s10648-016-9366-2

Perfetti, C. A. (1985). *Reading Ability*. New York, NY: Oxford University Press.

Poole, A. (2005). Gender differences in reading strategy use among ESL college students. *J. coll. Read. Learn.* 36, 7–20. doi: 10.1080/10790195.2005.10850177

Quinn, J. M. (2018). Differential identification of females and males with reading difficulties: a meta-analysis. *Read. Writ.* 31, 1039–1061. doi: 10.1007/s11145-018-9827-8

Quinn, J. M., and Wagner, R. K. (2015). Gender differences in reading impairment and in the identification of impaired readers: results from a large-scale study of at-risk readers. *J. Learn. Disabil.* 48, 433–445. doi: 10.1177/0022219413508323

Ramus, F. (2014). Neuroimaging sheds new light on the phonological deficit in dyslexia. *Trends Cogn. Sci.* 18, 274–275. doi: 10.1016/j.tics.2014.01.009

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., et al. (2004). Sex differences in developmental reading disability: new findings from 4 epidemiological studies. *J. Am. Med. Assoc.* 291, 2007–2012.

Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., and Escobar, M. D. (1990). Prevalence of reading disability in boys and girls: results of the connecticut longitudinal study. *J. Am. Med. Assoc.* 264, 998–1002. doi: 10.1001/jama.264.8.998

Sicular, T., Yue, X., Gustafsson, B., and Li, S. (2007). The urban-rural income gap and inequality in china. *Rev. Income and Wealth* 53, 93–126. doi: 10.1111/j.1475-4991.2007.00219.x

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia* 7, 37–46. doi: 10.1002/dys.185

Song, R. R., and Wu, H. R. (2008). Epidemic study on Chinese dyslexic children. *Maternal Child Health Care China* 23, 1505–1507.

Sparks, R. L. (1995). Examining the linguistic coding differences hypothesis to explain individual differences in foreign language learning. *Ann. Dyslexia* 45, 187–214. doi: 10.1007/BF02648218

Stanovich, K. E. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 21, 360–407. doi: 10.1598/rrq.21.4.1

Stevenson, H. W., Stigler, J. W., Lucker, G. W., Lee, S. Y., Hsu, C. C., and Kitamura, S. (1982). Reading disabilities: the case of Chinese, Japanese, and English. *Child Dev.* 53, 1164–1181. doi: 10.1111/j.1467-8624.1982.tb04154.x

Stoet, G., and Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: within-and across-nation assessment of 10 years of PISA data. *PLoS One* 8:e57988. doi: 10.1371/journal.pone.0057988

Tharp, R., and Gallimore, R. (1991). *The Instructional Conversation: Teaching and Learning in Social Activity*. Washington, DC: Office of Educational Research and Improvement.

Tong, X., Tong, X., and Mcbride-Chang, C. (2015). A tale of two writing systems: double dissociation and metalinguistic transfer between Chinese and English word reading among Hong Kong children. *J. Learn. Disabil.* 48, 130–145. doi: 10.1177/0022219413492854

Van Der Slik, F. W., Van Hout, R. W., and Schepens, J. J. (2015). The gender gap in second language acquisition: gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PLoS One* 10:e0142056. doi: 10.1371/journal.pone.0142056

Walberg, H. J., and Tsai, S. (1983). Matthew effects in education. *Am. Educ. Res. J.* 20, 359–373. doi: 10.3102/00028312020003359

Wang, J., Li, H., and Wang, D. (2018). Bridging the rural-urban literacy gap in China: a mediation analysis of family effects. *J. Res. Child. Educ.* 32, 119–134. doi: 10.1080/02568543.2017.1388308

Wang, M., Park, Y., and Lee, K. R. (2006). Korean-English biliteracy acquisition: cross-language phonological and orthographic transfer. *J. Educ. Psychol.* 98, 148–158. doi: 10.1037/0022-0663.98.1.148

Wang, X., and Tao, B. (1996). *Chinese Character Recognition Test Battery and Assessment Scale for Primary School Children*. Shanghai: Shanghai Education Press.

Woodcock, R. W. (1987). *Manual for the Woodcock Reading Mastery Test-Revised*. Circle Pines, MN: American Guidance Service.

Xuan, X., Xue, Y., Zhang, C., Luo, Y. H., Jiang, W., Qi, M. D., et al. (2019). Relationships among school socioeconomic status, teacher- student relationship, and middle school students' academic achievement in China: using the multilevel mediation model. *PLoS One* 14:e0213783. doi: 10.1371/journal.pone.0213783

Yang, M., Cooc, N., and Sheng, L. (2017). An investigation of cross-linguistic transfer between Chinese and English: a meta-analysis. *Asian Pac. J. Second Foreign Lang. Educ.* 2, 15–26.

You, H. L., Gaab, N., Wei, N., Cheng-lai, A., Wang, Z., Jian, J., et al. (2011). Neural deficits in second language reading: fMRI evidence from Chinese children with English reading impairment. *NeuroImage* 57, 760–770. doi: 10.1016/j.neuroimage.2010.12.003

Young, D. J. (1994). A comparison of student performance in Western Australian schools: rural and urban differences. *Austral. Educ. Res.* 21, 87–105. doi: 10.1007/bf03219569

Zhang, D., Li, X., and Xue, J. (2015). Education inequality between rural and urban areas of the people's republic of china, migrants' children education, and some implications. *Asian Dev. Rev.* 32, 32–45.

Zhang, H. C., and Wang, X. P. (1985). *Raven Standard Progressive Matrices: Chinese City Revision*. Beijing: The National Revision Collaborative Group.

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Zuo, P., Li, Z., and Wu, H. (2010). Study on the characteristics of Chinese developmental reading dyslexia among children. *Mod. Prev. Med.* 37, 3044–3046.

Universal and Specific Predictors of Chinese Children With Dyslexia – Exploring the Cognitive Deficits and Subtypes

Shuang Song[1,2†], Yuping Zhang[3†], Hua Shu[1,4*], Mengmeng Su[1,5] and Catherine McBride[6]*

[1] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China, [2] College of Teacher Education, Capital Normal University, Beijing, China, [3] Sichuan Research Center of Applied Psychology, Chengdu Medical College, Chengdu, China, [4] State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China, [5] Elementary Education College, Capital Normal University, Beijing, China, [6] Department of Psychology, The Chinese University of Hong Kong, Hong Kong, China

While previous studies have shown that the impact of phonological awareness (PA) and rapid automatized naming (RAN) on dyslexia depends on orthographic complexity in alphabetic languages, it remains unclear whether this relationship generalizes to the more complex orthography of Chinese. We investigated the predictive power of PA, RAN, and morphological awareness (MA) in dyslexia diagnosis status in a sample of 241 typically developing and 223 dyslexic Chinese-speaking children. Compared with the control group, children with dyslexia performed notably worse on character reading and all three cognitive measures. A logistic regression analysis showed that PA and RAN were both significant predictors, while MA also played a relatively important role for predicting dyslexia status in Chinese children. In the next step, we used multigroup analyses to test if these three cognitive predictors were of the same importance in predicting reading variance in different reading proficiency groups. And the results showed that the regression coefficient of MP is stronger for the control group than the dyslexia group, while the regression coefficient of PD tends to be stronger for the dyslexic group. Further cluster analysis identified four subtypes of dyslexia in this sample: a global deficit group, a phonological deficit group, a RAN deficit group, and a mild morphological deficit group. Our findings are largely consistent with previous studies of predictors of dyslexia, while uniquely demonstrating the differences in predictive power of these three cognitive variables on reading, as well as the unique contribution of MA in Chinese reading.

Keywords: dyslexia, phonology, morphology, subtypes, Chinese

INTRODUCTION

Developmental dyslexia is a specific disorder characterized by dysfluent or inaccurate word recognition that is not attributable to sensory deficit, insufficient education, or low IQ (Ramus et al., 2003). As researchers now widely support a multiple deficits model of Chinese reading difficulties (Ho et al., 2002; Zhou et al., 2014; Peng et al., 2017), one of the leading

questions researchers are focusing on is what are the important cognitive profiles of dyslexia in Chinese. In addition, given that there is no firm consensus from previous research (Ho et al., 2002; Liu et al., 2006; Wu et al., 2009), investigation of different subtypes of dyslexia is also a matter of interest. As it is relatively subjective to set a criteria for cognitive deficits in multiple-case analysis, cluster analyses based on a large sample can provide more reliable evidence about heterogeneous of dyslexia, as a data-driven method.

The multiple cognitive deficits model of reading difficulties proposed a multi-factorial etiology for this complex developmental disorder (Pennington, 2006). Accordingly, a number of cognitive causes have been put forward as of fundamental importance, ranging from cognitive anomalies possibly existing since long before formal education is received [deficits in phonological awareness (PA), or naming speed, for example] (Ziegler and Goswami, 2005; Furnes and Samuelsson, 2011), to deficits emerging with the acquisition of literacy and as consequences of impaired reading (such as orthographic deficit) (e.g., Ho et al., 2004; Xue et al., 2013). Therefore, in the present study, which aimed to examine the multi-deficit hypothesis in children from mainland China and to shed light on early markers for dyslexia, we only included assessments of the pre-literacy cognitive areas.

Research on predictors of reading abilities and disabilities has identified PA and rapid automatized naming (RAN, a measure of naming speed) as particularly strong indicators, both concurrently and longitudinally, even after statistically controlling for children's IQ (Wolf and Bowers, 1999; Ramus et al., 2003; White et al., 2006; Smythe et al., 2008; Furnes and Samuelsson, 2011). In comparison with normal controls, significant impairments in PA, and RAN have commonly been observed in dyslexic children (Pennington, 2006; Landerl et al., 2013). Although phonological skills have been suggested as an especially reliable predictor in predicting reading variation, evidence has shown that the strength of the relationship between the phonological deficit and reading varies with orthographic depth (Frost et al., 1987; Ziegler et al., 2010; Moll et al., 2014). Landerl et al. (2013) tested 1,138 typically developing children and 1,114 children with dyslexia across six orthographies with varying levels of consistency, and found that PA and RAN are both strong predictors of dyslexia diagnosis status. More interestingly, they reported that the impact of both cognitive domains is greater in complex orthographies (e.g., English) than in less complex orthographies (e.g., Finnish). As only alphabetic languages were involved in Landerl's study, it would be interesting to know whether the essence of how orthographic complexity influences reading can be generalized to Chinese, an even more complex orthography.

Over the past decade, research on predictors of dyslexia and poor reading in Chinese children has reported that both PA (McBride-Chang et al., 2011; Xue et al., 2013) and RAN (Ho et al., 2004; Pan et al., 2011) are strongly associated with children's reading variations, and dyslexic children were also observed as having significant deficits in PA and/or RAN (Shu et al., 2006; Chung et al., 2011). However, some other studies reported inconsistent findings: PA is not a significant predictor of

Chinese word reading among beginning readers after controlling for RAN, orthographic skills, and morphological awareness (MA) (Tong et al., 2009; Yeung et al., 2011). A recent meta-analysis of Chinese dyslexia found that the predictive power of PA for reading disability is not stable (Peng et al., 2017). The inconsistencies across these results may be partly due to the recruitment of relatively small samples of dyslexic children. Perhaps more importantly, there might be other significant factors correlated with learning to read Chinese, such as MA.

The morpheme, as the smallest meaningful unit, provides basic semantic information within a language. MA refers to the ability to manipulate morphemes and employ word information rules in one language (Shu et al., 2006). One of the most prominent characteristics of Chinese is that the language contains a large number of homophones, with Mandarin having an average of five homophones corresponding to each syllable, taking tones into consideration. For example, the syllable/qing1/can represent more than six characters with different meanings [e.g., 青(blue), 清(clear), 蜻(dragonfly), 轻(relax), 氢(hydrogen), 卿(a minister or a high official in ancient times)]. This phenomenon leads to ambiguity when only sound is used to distinguish words in Chinese. Thus, to become a successful reader one needs to be able to distinguish the meanings of words that sound identical (Shu et al., 2006; Tong et al., 2009). A series of Chinese studies have demonstrated that MA is associated with literacy skills and reading disability (McBride-Chang et al., 2003; Li et al., 2012; Liu et al., 2013). MA has been found to be both a concurrent and longitudinal predictor of reading in typically developing children (Lei et al., 2011; Liu et al., 2013). Moreover, MA has also been found to be one of the best factors that distinguishing dyslexia Chinese children from their age-matched controls (Shu et al., 2006; Zhou et al., 2014; Tong et al., 2017)., and children with severe reading deficits generally show more sever deficits in MA (Peng et al., 2017). Given the aforementioned property of Chinese, it is not hard to understand why MA has been thoroughly accepted as both a strong concurrent and a longitudinal predictor of Chinese literacy skills (Chen et al., 2009; Yeung et al., 2011; Pan et al., 2015). Therefore, in the present study, a morphological production task (suitable for older children) was also included to tap into children's MA (e.g., Shu et al., 2006).

Another set of studies devoted to identifying subgroups of dyslexia has revealed that people with dyslexia who have a deficit in PA constitute the most commonly identified subgroup across languages (Ho et al., 2004; White et al., 2006), while some studies have found that children with dyslexia also have difficulty with rapid naming (Katzir et al., 2008; Jednoróg et al., 2014). Compared with multiple-case analysis, cluster analyses avoid debate over the criteria for cognitive deficits by adopting a data-driven method for the classification of subgroups (Heim et al., 2008; Jednoróg et al., 2014).

With different subtypes reported across studies, it is possible that the dyslexic population is heterogeneous, with varying degrees of impairment in different cognitive skills (White et al., 2006; Ho et al., 2007). Unlike alphabetic languages, Chinese is characterized by its morphosyllabic writing system where 90% of the Chinese characters consist of two components: the semantic radical gives a clue to meaning and the phonetic

gives a clue to pronunciation. In addition, new words in the Chinese language are made up of novel combinations of existing syllables, and not through the coining of new syllables. Dyslexia subtypes in Chinese samples have therefore been found to show characteristics specific to the language itself (Ho et al., 2002; Shu et al., 2006). For example, Wu et al. (2009) report observing that, in a group of 75 Chinese children with dyslexia, the largest proportion (96%) exhibited deficits in MA, compared with 53% with deficits in PA and 45% with deficits in RAN. Until now, there has been no firm consensus on which subtypes of dyslexia exist among Chinese speakers, and this necessitates further exploration.

To summarize, the current study addressed one research question from two dimensions. First, with the same case-control design and data-driven logistic regression analyses as used as in Landerl et al. (2013) research, the current study aimed to test whether these preliteracy cognitive areas of PA, RAN, and MA are important predictors in a large sample of Chinese children ($n = 464$), and whether these cognitive variables are contributing differently in predicting reading between dyslexia and normal developing children. The second goal was to identify specific dyslexic subtypes in the Chinese language based on an investigation of the three above-mentioned cognitive domains.

MATERIALS AND METHODS

Participants

Ethical approval for the present study was obtained from the Institutional Review Board at Beijing Normal University. In total, 223 individuals with dyslexia and 241 typically developing controls participated in the present study. All were children born in Beijing, China, and were native Mandarin speakers, with normal IQ and no reported mental, physical, or sensory difficulties.

Children diagnosed with dyslexia were recruited from eleven elementary schools in Beijing, attended by a total of about 3,600 children aged between 9 and 11 years. Children with dyslexia were identified and selected using the following procedure: (1) as recommended by their Chinese teachers, the lowest 20% school reading academic performance children (total $n = 708$) in each class were invited to participate the screening test for dyslexia; (2) Of these children, those who scored at least 1.5 SD below their respective grade mean on the Chinese character reading (CR) (Xue et al., 2013) were included. This threshold was based on previously used criteria (Shu et al., 2006); (3) Those with either performance IQ or full-scale IQ scores lower than 85 on the Wechsler Intelligence Scale for Children (C-WISC; Gong and Cai, 1993) were excluded. The remaining children were identified as having dyslexia, and thus participated in data collection (see section "Measures" below); and (4) On the basis of parental reports on a rating scale for attention and hyperactivity behavior, data from children identified as having attention deficit hyperactivity disorder (Swanson, Nolan, and Pelham –IV Teacher and Parent 18-Item Rating Scale, Swanson et al., 2001) were additionally excluded from analyses.

This resulted in a sample of 201 dyslexic children (mean age = 130.70 ± 17.07 months; 155 boys). Using the same criteria

as described above, a further 22 children with dyslexia (mean age = 124.36 ± 4.29 months; 17 boys) and with a normal IQ score of ≥25th percentile on the Raven's Standard Progressive Matrices were identified from an ongoing longitudinal cohort study (Lei et al., 2011), resulting in a total sample of 223 children with dyslexia.

The control group consisted of 241 children, all from the aforementioned longitudinal cohort (mean age = 125.27 ± 3.55 months; 128 boys). The inclusion criteria were: (1) a score no more than 1 SD below the grade mean on the same character reading test as mentioned above, and (2) a normal IQ score of ≥25th percentile on the Raven test. Subtests of C-WISC test were administered for only 131 children in control group.

Measures

Chinese Character Reading

This task was used to measure children's untimed reading accuracy (Lei et al., 2011). The CR task consists of a list of 150 single Chinese characters, which the children were asked to name; self-corrections and guessing were allowed. All of these characters are expected to have been learned by grade six in Beijing (Shu et al., 2003). The final score was the total number of characters that a child correctly named. Cronbach's alpha for the CR task is 0.94.

Phoneme Deletion

The PD task was used to measure children's PA (Pan et al., 2015). Participants were required to delete a target phoneme from a monosyllabic Chinese word (e.g., "Say/shu1/without the/sh/"). The target phoneme for each item was the first, middle, or final phoneme. The test consisted of two practice items and 26 experimental items, and the final score was the number of correctly answered experimental items. Cronbach's alpha for the PD task is 0.83.

Rapid Automatized Naming of Digits

The RAN task consisted of a 5 × 10 matrix of digits that children were required to name as quickly and accurately as possible (Pan et al., 2011). This task was administered twice, and the mean total naming time across the two trials was taken as the final score. The test–retest reliability of the RAN task is 0.92.

Morphological Production

The MP task has been widely used in previous studies to measure Chinese children's MA (Shu et al., 2006). During the test, participants were orally presented with 15 two-character compound words with one of the morphemes highlighted as the target (e.g., in the word 阳光/yang2guang1/, meaning sunshine, the target morpheme was 光/guang1/). They were then required to orally produce two new words containing the same Chinese character as the target morpheme. In one of these cases, the morpheme represented by this character in the new word should be the same as the target morpheme (e.g., a one-point answer in the above case was 月光/yue4guang1/, meaning moonlight). In the other case, the morpheme represented by this character should be different from the original target morpheme

(e.g., 光滑/guang1hua2/, meaning smooth, contains a homograph morpheme with a different meaning and would be scored as one point). Answers not produced according to the guidelines were scored as zero. The final score was the number of correct words given during the task, with a maximum score of 30. Cronbach's alpha for the MP task is 0.80.

Statistical Analyses

Raw scores on all measures were converted to grade-specific z-scores according to a previous large-scale reading study (Xue et al., 2013); these z-scores were entered into all subsequent analyses. The grade-specific z-scores for the RAN digits test were multiplied by −1; thus, higher scores represented better performance, as for the other measures. Deficits in the cognitive domains of PA, MA, and RAN were defined as a grade-specific z-score below −1 SD.

Predictive Analysis

SPSS 20.0 was used to conduct a logistic regression analysis (Peng et al., 2002). The Hosmer–Lemeshow test was used to assess the goodness of fit of the model. Scores on each of the three cognitive skill tests, namely PD, RAN digits, and MA, were introduced as predictors of dyslexic status, in accordance with the following model:

$$P(\text{diagnosis with dyslexia}) = 1/\{1 + \exp[-(\beta_0 + \beta_1 \times (PD)_i$$
$$+\beta_2 \times (RAN)_i + \beta_3 \times (MA)_i + \varepsilon_i)]\}$$

Multi-Group Analysis

Linear regression analyses were used to examine the predictability of PD, RAN, and MA on children's reading ability separately for dyslexics and normal controls. To test the equality across the control and dyslexic groups, multi-group analyses were conducted to test the possible different effects of these cognitive skills on the reading performance. The chi-square difference test comparing constrained models and freely estimated models was used to evaluate the model.

Subtype Analysis

In order to explore the subtypes within the dyslexic sample, cluster analysis (Rousseeuw, 1987) was used, as follows: (1) Hierarchical clustering was first used to determine the number of subgroups based on the three cognitive measures of PD, RAN, and MA, with larger changes in agglomeration coefficient representing a better-fitting number of clusters (Jurowski and Reich, 2000); (2) Subsequently, the k-means technique was applied to identify the final clusters. In the first step, between-groups linkage was used to combine dyslexic children into clusters during hierarchical clustering. Each squared Euclidean distance between two data points was calculated as a measure of similarity of two participants. The best cluster solution was selected by visual inspection of the agglomeration coefficients. The Average Silhouette Width was also used to suggest the "best number" of clusters. In the second step, k-means clustering was used to maximize cluster homogeneity and the number of clusters was decided from hierarchical procedure. this approach revealed the characteristics of different patterns of dyslexia. We then used

ANOVA and Turkey's *post hoc* test to test the differences among subgroups of dyslexia and control group separately for all three cognitive skills and the reading measures. This approach revealed the characteristics of different patterns of dyslexia.

RESULTS

Descriptive statistics for raw scores and grade-specific z-scores for reading and the three cognitive tests are presented in **Table 1**. Compared with control group children, children with dyslexia performed notably worse on the character reading task (CR: 0.33 vs. − 2.45). The additional clear group differences in scores on the three cognitive skills tasks suggested that these three profiles could be useful in distinguishing children with and without dyslexia. Correlations among reading and cognitive measures are shown in **Supplementary Table S1**, with all measures significantly intercorrelated.

Next, PD, RAN digits, and MP test scores were introduced as predictive variables in a logistic regression model. The corresponding odds ratios (OR) and coefficient estimates (ln OR) derived from the Wald statistic are presented in **Table 2** (Model 1). As expected, PD, RAN digits, and MP scores were all reliable predictors of dyslexia status. Participants with lower PD, RAN digits, or MP scores were at an increasing risk of having being diagnosed with dyslexia (PD: OR = 0.52, 95% CI [0.41, 0.67]; RAN digits: OR = 0.45, 95% CI [0.33, 0.60]; MP: OR = 0.25, 95% CI [0.17, 0.35]). As shown in **Figure 1**, the association was stronger in the case of MP than in the case of either PD or RAN digits. A logistic regression analysis controlling for gender, age, block design, and similarities was also conducted for children to whom the C-WISC test had been administered ($n = 343$), and similar predictive patterns were observed for the three core cognitive measures (Model 2 in **Table 2**).

In the next step, we used regression analyses to test if these three cognitive predictors were of the same importance in predicting reading. As can be seen in **Table 3**, the results differed for the dyslexia and control groups. For the control children, both RAN (β = 0.23, $P < 0.001$) and MP (β = 0.31, $P < 0.001$) were strong predictors, while PD was not (β = 0.07, $P = 0.249$), indicating that better performance on RAN and/or MP was associated with better performance in reading. For the dyslexic children, PD (β = 0.25, $P = 0.001$) and RAN (β = 0.13, $P = 0.048$) were two strong predictors, while MP was not (β = 0.09, $P = 0.206$), indicating that poor PA and/or slower RAN, but especially PA, was usually combined with poor performance in reading. The change in the χ^2 values when each predictor was constrained to be equal for these two groups during multi-group analyses was also included in **Table 3**. There was no significant change when the prediction strength between RAN and CR was constrained to be equal. However, the regression coefficient of MP is stronger ($\chi^2 = 4.026$, $p = 0.045$) for the control group than the dyslexic group. Moreover, the regression coefficient of PD tends to be stronger ($\chi^2 = 3.246$, $p = 0.072$) for the dyslexic group.

Subsequently, cluster analysis was carried out to explore subtypes in the dyslexic sample. Change in the agglomeration coefficients suggested that a two-cluster or four-cluster model

TABLE 1 | Descriptive statistics for all reading measures for control and dyslexia groups.

	Raw mean (SD)	Z Mean (SD)	Min	Max	Skewness	Kurtosis
Controls (n = 241, 53% boys)						
PD	18.85 (1.66)	0.08 (0.92)	−3.74	0.73	−1.81	3.00
RAN	15.91 (3.55)	0.13 (0.92)	−4.42	2.16	−1.13	2.84
MP	22.97 (3.42)	0.18 (0.88)	−2.38	1.99	−0.41	−0.29
CR	118.48 (9.93)	0.33 (0.69)	−0.94	1.75	0.08	−0.89
Dyslexics (n = 223, 77% boys)						
PD	12.27 (5.19)	−1.70 (1.54)	−5.29	2.47	−0.20	−0.31
RAN	20.82 (4.29)	−1.12 (1.22)	−5.89	1.34	−0.96	1.45
MP	16.93 (4.05)	−1.47 (1.09)	−4.73	0.99	−0.37	0.00
CR	85.14 (13.39)	−2.45 (1.03)	−8.88	−1.52	−2.42	8.58

PD, Phoneme deletion; RAN, rapid automatized naming digits; MP, morphological production; CR, Chinese character reading.

TABLE 2 | The logistic regression model for predicting dyslexia.

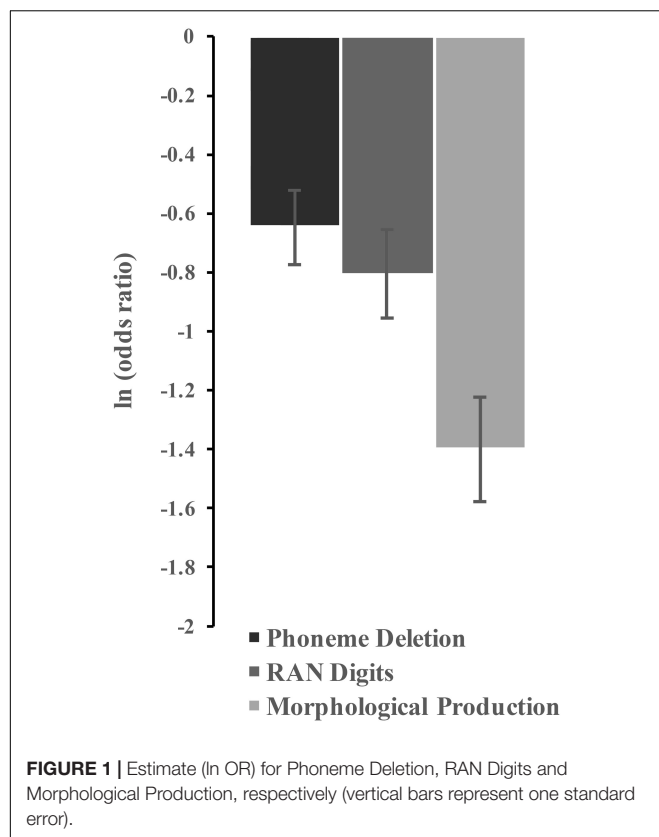|	Estimate	S.E.	Odds Ratio (OR)	p-Value	95% C.I. for OR	
					Present study	Landerl et al., 2013
Model 1[a] (n = 464)						
PD	−0.65	0.13	0.52	<0.001	[0.41, 0.67]	[0.31, 0.41]
RAN	−0.81	0.15	0.45	<0.001	[0.33, 0.60]	[0.31, 0.41]
MP	−1.40	0.18	0.25	<0.001	[0.17, 0.35]	–
Cox & Snell R Square					0.511	
Model 2[b] (n = 343): 212 dyslexics and 131 controls						
Gender	0.83	0.45	2.30	0.065	[0.95, 5.58]	
Age	0.06	0.02	1.06	0.002	[1.02, 1.09]	
Block design[c]	−0.39	0.09	0.68	<0.001	[0.57, 0.81]	
Similarities[d]	−0.18	0.11	0.84	0.104	[0.67, 1.04]	
PD	−0.68	0.18	0.51	<0.001	[0.35, 0.72]	
RAN	−0.66	0.21	0.52	0.002	[0.34, 0.79]	
MP	−1.20	0.25	0.30	<0.001	[0.18, 0.49]	
Cox & Snell R Square					0.578	

PD, Phoneme deletion; RAN, rapid automatized naming digits; MP, morphological production; CR, Chinese character reading. a, full sample included. b, only children with WISC-R scores included. c, d, in WISC-R.

would best capture the data. When referring to the average silhouette width, the value for two-cluster solution and which for four-cluster solution are both local maximum values. However, the two-cluster solution only showed an overall assessment of high or low performance and did not reveal the differences among subgroups. In order to understand the features of different deficit patterns more clearly, the four-cluster solution was applied. To validate the results of the four-group cluster analysis, discriminant analysis was used. As suggested by the leave-one-out classification during discriminant analysis, 96.4% of original grouped cases were correctly classified.

Figure 2 presents the characteristics of the deficit patterns for the four groups, which were labeled as a global deficit group (Group 1), a phonological deficit group (Group 2), a RAN deficit group (Group 3), and a mild morphological deficit group (Group 4). Participants in the global deficit group exhibited difficulty in all three cognitive domains (**Table 4**), with grade-specific z-scores all below −1.5. Specifically, participants in this group performed particularly poorly on PD ($z = -3.77$) and MP ($z = -2.59$).

The phonological deficit group obtained the lowest scores on PD ($z = -2.13$) compared to the performance of children in other groups, with the exception of the global deficit group. Children in the phonological deficit group also exhibited poor MA ($z = -1.39$), while their performance in RAN was only moderately impaired ($z = -0.75$), although still lower than that of the control group. Along with poor PA ($z = -1.18$), the RAN deficit group scored the lowest of all four subgroups on RAN ($z = -2.71$). However, their performance in MP was no worse than that of children in either the phonological deficit group or the mild morphological deficit group. Children in the mild morphological deficit group exhibited difficulty only in MP ($z = -1.07$), with almost normal performance on both the PD and RAN tasks.

Participants in the four subgroups also showed heterogeneity in their CR task scores (**Table 4**). Children in the mild morphological deficit group only scored lower than those in the control group in CR (cohen's $d = -3.87$, $p < 0.001$), and performed better than the other three subgroups (Group 1: cohen's $d = 0.89$; Group 2: cohen's $d = 0.76$; Group 3: cohen's

FIGURE 1 | Estimate (ln OR) for Phoneme Deletion, RAN Digits and Morphological Production, respectively (vertical bars represent one standard error).

$d = 0.82$; ps < 0.001). With performance significantly worse than those in the control and mild morphological deficit groups, children in the other three dyslexia subgroups performed comparably on the Chinese character reading (Group 3 vs. Group 2: cohen's $d = 0.08$; Group 2 vs. Group 1: cohen's $d = 0.33$; Group 3 vs. Group 1: cohen's $d = 0.41$; ps > 0.05).

DISCUSSION

With a relatively large sample of 223 dyslexic children and 241 typically developing children, the present study investigated to what extent various cognitive variables predicted children's diagnostic status, and the differences in predictive power of these variables on reading. In addition, characteristics of different dyslexia subtypes in Chinese had also been examined. Similarly to the findings for alphabetic languages (Landerl et al., 2013; Moll et al., 2014), our results also showed that deficits in PA and RAN tests were both strong predictors of dyslexia status. However, a deficit in MA was the best predictor of Chinese dyslexia. As predicting individual variances in children's reading, MA showed stronger predictability in the control group, while PA tends to be a better predictor in the dyslexia group, and RAN was of equal importance for both groups. Furthermore, we identified four subgroups of different deficit patterns and suggested that more severe dyslexics were particularly more impaired in these three cognitive domains. However, it is worth noting that all four subgroups

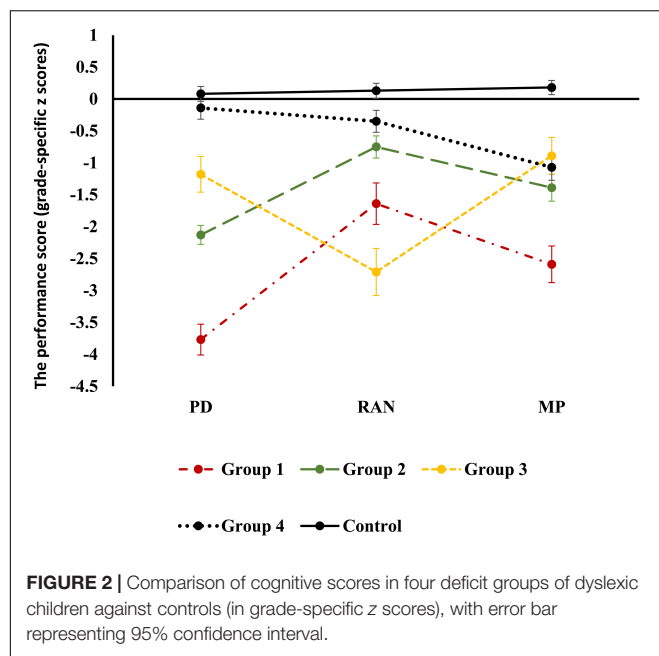**TABLE 3 |** Standardized coefficients in linear regression and multi-group analyses.

	CR					
	Dyslexic group		Control group		Multi-group analyses	
	β	*p*-Value	β	*p*-Value	χ²	*p*-Value
PD	0.252	0.001	0.069	0.249	3.246	0.072
RAN	0.131	0.048	0.233	<0.001	0.806	0.369
MP	0.088	0.206	0.313	<0.001	4.026	0.045

PD, Phoneme deletion; RAN, rapid automatized naming digits; MP, morphological production; CR, Chinese character reading.

exhibited moderate to severe deficits in MA and that the observation of heterogeneity in Chinese dyslexia is consistent with previous findings (Ho et al., 2004; Shu et al., 2006; Wu et al., 2009).

Both PA and RAN in the present study, either with the morphological factor included (**Table 2**) or excluded (**Supplementary Table S2**), significantly distinguished dyslexia status. What's more interesting is that these cognitive skills differed in predicting individual differences for Chinese children with and without dyslexia. Compared with the control group, PA showed a stronger prediction on reading in the dyslexia group. One possible explanation could be that PA developmentally serves as a base for children learning to read (Hong et al., 2018), and it may contribute to reading acquisition through its impact on MA (Pan et al., 2015). As a result, phonology would be more influential for the lower-proficiency Chinese readers. For children at a higher level of reading achievement, variance in reading accuracy is reduced, and similar variance reductions can also be found in PA. Thus, the role of PA may become less salient with reading experience. Similarly, some previous studies showed for readers in five alphabetic languages that the predictive power of PA became weaker as mastering of orthography–phonology correspondence became easier (such as in highly transparent languages) (Ziegler et al., 2010; Furnes and Samuelsson, 2011). Together with the results from the present study, these findings underscore the universality of the importance of PA in reading, and shed light on the need to pay attention to different proficiencies for understanding the phonology–reading relationship.

In line with previous findings in alphabetic languages and Chinese (Kirby et al., 2010; Landerl et al., 2013; Zhou et al., 2014; Peng et al., 2017), RAN significantly predicted children's diagnostic status and individual variances in reading. Furthermore, our multi-group analyses revealed that the predictability of RAN on reading is comparable between the dyslexic and control groups, suggesting a dominant role of this skill in Chinese children across different reading proficiencies. Establishing fluency in reading, involving automatic sequencing of Chinese characters, is of particular importance to become a skilled reader. The rapid number naming task tapped this ability across reading levels, such that those who were faster and more effective at the orthographic–phonology accessing in one domain (i.e., numbers), also tend to be better readers in the domain of character reading.

FIGURE 2 | Comparison of cognitive scores in four deficit groups of dyslexic children against controls (in grade-specific *z* scores), with error bar representing 95% confidence interval.

More importantly, our study suggested that MA, in addition to PA and RAN, appears to be an even more important cognitive construct in Chinese dyslexia. This was in line with a series of previous studies of Chinese reading development and impairment (McBride-Chang et al., 2003; Chen et al., 2009; Tong et al., 2017). It is interesting to note that MA exhibited greater predictive power than did phonological processing skills in the present study. This is partially due to the properties of the Chinese writing system and the nature of the process of learning to read Chinese. Categorized as a morphosyllabic language, Chinese is relatively semantically transparent; therefore, mastering the meaning of the character or morpheme is vitally important in learning to read, especially in new characters and word learning. Moreover, Chinese contains many homophones, so knowing the meaning of morphemes may greatly help children to distinguish these homophones and thus improve their reading ability (e.g., McBride-Chang et al., 2003; Shu et al., 2006). Thus, it is no wonder that MA was found to be more important for the skilled readers, the controls, in the present study. In Chinese, semantic

information is reflected by the morphological properties of words. Therefore, MA is key for Chinese character recognition, a widely used measure in dyslexia diagnosis (Li et al., 2012), as it helps children to distinguish the meanings of different characters.

Four subgroups of children with dyslexia were identified in the present study and these subtype characteristics also supported our findings in the predictive analyses. When morphological skills were comparable (Groups 2 and 3 vs. Group 4), poor PA and/or slower naming speed would block one's possibility of becoming a better reader, suggesting phonology as an essential factor in deficient readers. On the other hand, with poor PA and slower naming speed (Group 1 vs. Group 3), additional inferior skills in morphology only slightly impair children's reading and indicate a diminished function of MA. Along with the aforementioned findings, it is clear that morphological skills exhibited a more important role in skilled readers as compared with deficient readers. This is because for a beginning reader, most of the characters to be learned are simple characters, which are more or less directly meaningful (Shu et al., 2003). Thus, correspondence among orthography, phonology, and meaning are relatively transparent. For a skilled reader, however, due to the characteristics of the Chinese language, the ability to distinguish between homophones and homographs becomes increasingly important. As a result, the importance of MA gradually emerges. To summarize briefly, as the importance of different cognitive skills varies, it is necessary to provide differentiated educational strategies for children of different reading proficiency.

In addition, children in the first three groups (the global deficit group, phonological deficit group, and RAN deficit group) all showed moderate to severe deficits in PA. Thus, phonological deficits represented one dominant characteristic across the whole sample (68.2%). This is similar to what has been found in previous multiple-case studies, in which phonological deficits have been found to emerge in the vast majority of participants with dyslexia (Ramus et al., 2003; White et al., 2006), but differs from the findings of Ho et al. (2002, 2004), who observe that only 15.3 to 29.3% of children with dyslexia in Hong Kong exhibit deficits in phonology. This disparity may be partially explained by the different forms of instruction used in the relevant education systems as well as differences between Mandarin and Cantonese. In Hong Kong, a whole-word and drilling approach is used,

TABLE 4 | Means (SD) grade-specific Z-scores of the classification measures of control and four deficit groups.

		Dyslexia subgroups				
	Control (241)	Group 1 (48)	Group 2 (69)	Group 3 (35)	Group 4 (71)	*F* value
PD	0.08 (0.92)ₐ	−3.77 (0.85)ₑ	−2.13 (0.63)_d	−1.18 (0.84)_c	−0.14 (0.76)_b	266.83***
RAN	0.13 (0.92)ₐ	−1.64 (1.14)_d	−0.75 (0.74)_c	−2.71 (1.11)ₑ	−0.35 (0.74)_b	100.63***
MP	0.18 (0.88)ₐ	−2.59 (1.02)_c	−1.39 (0.90)_b	−0.89 (0.86)_b	−1.07 (0.88)_b	124.39***
CR	0.33 (0.69)ₐ	−2.97 (1.51)_c	−2.55 (0.97)_c	−2.48 (0.74)_c	−1.97 (0.48)_b	332.88***

*PD, Phoneme deletion; RAN, rapid automatized naming digits; MP, morphological production; CR, Chinese character reading. Group 1, global deficit; Group 2, phonological deficit; Group 3, RAN deficit; Group 4, mild morphological deficit. ***p < 0.001. Means in the same row that do not share subscripts differ at p < 0.05 on Turkey's post hoc test.*

in which children must retrieve the pronunciation of characters by rote (Ho et al., 2004), while children in mainland China learn to assemble the sounds of characters through the pinyin phonetic system. The latter approach emphasizes the importance of phonological processing in reading.

There were some limitations to the present study. First, although we intended to investigate the predictability of preliterate cognitive skills, it might have been also interesting to have included some postliterate cognitive domains (orthographic awareness for example). Second, the present findings are all based on a cross-sectional design. With this study, in addition to several previous studies of Chinese reading impairment, we are beginning to get a better understand of the essential role of these cognitive skills. However, more longitudinal studies should be introduced to examine the contribution of these skills over time. Moreover, in future studies, reading impairments in other literacy domains should be considered, such as in reading fluency, spelling, and reading comprehension.

Despite these limitations, the current findings suggest some important conclusions about Chinese dyslexia. First, our findings highlight the important effects of MA, in addition to the effects of PA and the data from RAN tests, in understanding developmental dyslexia. The predictive power of morphological processing in explaining dyslexia in Chinese speakers suggests the necessity of acquiring knowledge about morphemes, which may help one to become a more skilled reader. Second, the contribution of the three cognitive skills differed across children's reading proficiency, indicating that differential educational strategies should be taken into consideration in teaching (more attention should be addressed on phonological rules for beginning readers and/or slow learners, for example). Finally, our findings indicate that dyslexic Chinese children are heterogeneous, and the majority of children exhibited double or multiple deficits, which provided additional support for the multiple-deficit hypothesis for Chinese developmental dyslexia.

DATA AVAILABILITY STATEMENT

The datasets for this manuscript are not publicly available because: the present data belongs to an ongoing longitudinal study and most of the data are still in processing and need to be protected. Requests to access the datasets should be directed to HS, shuh@bnu.edu.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Beijing Normal University Imaging Center for Brain Research and the State Key Laboratory of Cognitive Neuroscience and Learning. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

HS, YZ, and CM contributed to conception and design of the study. YZ organized the database. SS and YZ performed the statistical analysis and wrote the manuscript. All authors contributed to manuscript revision and read and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02904/full#supplementary-material

REFERENCES

Chen, X., Hao, M., Geva, E., Zhu, J., and Shu, H. (2009). The role of compound awareness in Chinese children's vocabulary acquisition and character reading. *Read. Writ.* 22, 615–631. doi: 10.1007/s11145-008-9127-9

Chung, K., Ho, C., Chan, D., Tsang, S.-M., and Lee, S.-H. (2011). Cognitive skills and literacy performance of Chinese adolescents with and without dyslexia. *Read. Writ.* 24, 835–859. doi: 10.1007/s11145-010-9227-1

Frost, R., Katz, L., and Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: a multilingual comparison. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 104–115. doi: 10.1037//0096-1523.13.1.104

Furnes, B., and Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: results from a cross-linguistic longitudinal study. *Learn. Individ. Differ.* 21, 85–95. doi: 10.1016/j.lindif.2010.10.005

Gong, Y., and Cai, T. (1993). *Wechsler Intelligence Scale for Children, Chinese Revision (C-WISC)*. China: Map Press Hunan.

Heim, S., Tschierse, J., Amunts, K., Wilms, M., Vossel, S., Willmes, K., et al. (2008). Cognitive subtypes of dyslexia. *Acta Neurobiol. Exp.* 68, 73–82.

Ho, C., Chan, D., Chung, K., Lee, S., and Tsang, S. (2007). In search of subtypes of Chinese developmental dyslexia. *J. Exp. Child Psychol.* 97, 61–83. doi: 10.1016/j.jecp.2007.01.002

Ho, C., Chan, D., Lee, S., Tsang, S., and Luan, V. (2004). Cognitive profiling and preliminary subtyping in Chinese developmental dyslexia. *Cognition* 91, 43–75. doi: 10.1016/S0010-0277(03)00163-X

Ho, C., Chan, D., Tsang, S., and Lee, S. (2002). The cognitive profile and multiple-deficit hypothesis in Chinese developmental dyslexia. *Dev. Psychol.* 38, 543–553. doi: 10.1037//0012-1649.38.4.543

Hong, T., Shuai, L., Frost, S. J., Landi, N., Pugh, K. R., and Shu, H. (2018). Cortical responses to chinese phonemes in preschoolers predict their literacy skills at school age. *Dev. Neuropsychol.* 43, 356–369. doi: 10.1080/87565641.2018.1439946

Jednoróg, K., Gawron, N., Marchewka, A., Heim, S., and Grabowska, A. (2014). Cognitive subtypes of dyslexia are characterized by distinct patterns of grey matter volume. *Brain Struct. Funct.* 219, 1697–1707. doi: 10.1007/s00429-013-0595-6

Jurowski, C., and Reich, A. Z. (2000). An explanation and illustration of cluster analysis for identifying hospitality market segments. *J. Hosp. Tour. Res.* 24, 67–91. doi: 10.1177/109634800002400105

Katzir, T., Kim, Y. S., Wolf, M., Morris, R., and Lovett, M. W. (2008). The varieties of pathways to dysfluent reading: comparing subtypes of children with dyslexia at letter, word, and connected text levels of reading. *J. Learn. Disabil.* 41, 47–66. doi: 10.1177/0022219407311325

Kirby, J. R., Georgiou, G. K., Martinussen, R., and Parrila, R. (2010). Naming speed and reading: from prediction to instruction. *Read. Res. Q.* 45, 341–362. doi: 10.1598/RRQ.45.3.4

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P., Lohvansuu, K., et al. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *J. Child Psychol. Psychiatry* 54, 686–694. doi: 10.1111/jcpp.12029

Lei, L., Pan, J., Liu, H. Y., McBride-Chang, C., Li, H., Zhang, Y. P., et al. (2011). Developmental trajectories of reading development and impairment from ages 3 to 8 years in Chinese children. *J. Child Psychol. Psychiatry* 52, 212–220. doi: 10.1111/j.1469-7610.2010.02311.x

Li, H., Shu, H., McBride-Chang, C., Liu, H., and Peng, H. (2012). Chinese children's character recognition: visuo-orthographic, phonological processing and morphological skills. *J. Res. Read.* 35, 287–307. doi: 10.1111/j.1467-9817.2010.01460.x

Liu, P. D., McBride-Chang, C., Wong, T. T.-Y., Shu, H., and Wong, A. M.-Y. (2013). Morphological awareness in Chinese: unique associations of homophone awareness and lexical compounding to word reading and vocabulary knowledge in Chinese children. *Appl. Psycholinguist.* 34, 755–775. doi: 10.1017/s014271641200001x

Liu, W. L., Liu, X. P., and Zhang, J. Q. (2006). A preliminary study subtypes of Chinese developmental dyslexia. *Acta Psychologica Sinica* 38, 681–693. doi: 10.1016/S0379-4172(06)60092-9

McBride-Chang, C., Lam, F., Lam, C., Chan, B., Fong, C. Y.-C., Wong, T. T.-Y., et al. (2011). Early predictors of dyslexia in Chinese children: familial history of dyslexia, language delay, and cognitive profiles. *J. Child Psychol. Psychiatry* 52, 204–211. doi: 10.1111/j.1469-7610.2010.02299.x

McBride-Chang, C., Shu, H., Zhou, A., Wat, C. P., and Wagner, R. K. (2003). Morphological awareness uniquely predicts young children's Chinese character recognition. *J. Educ. Psychol.* 95, 743–751. doi: 10.1037/0022-0663.95.4.743

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., et al. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learn. Instr.* 29, 65–77. doi: 10.1016/j.learninstruc.2013.09.003

Pan, J., McBride-Chang, C., Shu, H., Liu, H., Zhang, Y., and Li, H. (2011). What is in the naming? A 5-year longitudinal study of early rapid naming and phonological sensitivity in relation to subsequent reading skills in both native Chinese and English as a second language. *J. Educ. Psychol.* 103, 897–908. doi: 10.1037/a0024344

Pan, J., Song, S., Su, M., McBride, C., Liu, H., Zhang, Y., et al. (2015). On the relationship between phonological awareness, morphological awareness and Chinese literacy skills: evidence from an 8-year longitudinal study. *Dev. Sci.* 19, 982–991. doi: 10.1111/desc.12356

Peng, C.-Y., Lee, K., and Ingersoll, G. (2002). An introduction to logistic regression analysis and reporting. *J. Educ. Res.* 96, 3–14. doi: 10.1080/0022067020959 8786

Peng, P., Wang, C., Tao, S., and Sun, C. (2017). The deficit profiles of chinese children with reading difficulties: a meta-analysis. *Educ. Psychol. Rev.* 29, 513–564. doi: 10.1007/s10648-016-9366-2

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition* 101, 385–413. doi: 10.1016/j.cognition.2006.04.008

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126, 841–865. doi: 10.1093/brain/awg076

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65. doi: 10.1016/0377-0427(87)90125-7

Shu, H., Chen, X., Anderson, R. C., Wu, N., and Xuan, Y. (2003). Properties of school Chinese: implications for learning to read. *Child Dev.* 74, 27–47. doi: 10.1111/1467-8624.00519

Shu, H., McBride-Chang, C., Wu, S., and Liu, H. (2006). Understanding chinese developmental dyslexia: morphological awareness as a core cognitive construct. *J. Educ. Psychol.* 98, 122–133. doi: 10.1037/0022-0663.98.1.122

Smythe, I., Everatt, J., Al-Menaye, N., He, X., Capellini, S., Gyarmathy, E., et al. (2008). Predictors of word-level literacy amongst Grade 3 children in five diverse languages. *Dyslexia* 14, 170–187. doi: 10.1002/dys.369

Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., et al. (2001). Clinical relevance of the primary findings of the mta: success rates based on severity of ADHD and odd symptoms at the end of treatment. *J. Am. Acad. Child Adolesc. Psychiatry* 40, 168–179. doi: 10.1097/00004583-200102000-00011

Tong, X., Mcbride, C., Lo, J. C. M., and Shu, H. (2017). A three-year longitudinal study of reading and spelling difficulty in Chinese developmental dyslexia: the matter of morphological awareness. *Dyslexia* 23, 372–386. doi: 10.1002/dys.1564

Tong, X., McBride-Chang, C., Shu, H., and Wong, A. M. (2009). Morphological awareness, orthographic knowledge, and spelling errors: keys to understanding early Chinese literacy acquisition. *Sci. Stud. Read.* 13, 426–452. doi: 10.1080/10888430903162910

White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., et al. (2006). The role of sensorimotor impairments in dyslexia: a multiple case study of dyslexic children. *Dev. Sci.* 9, 237–255. doi: 10.1111/j.1467-7687.2006.00483.x

Wolf, M., and Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *J. Educ. Psychol.* 91:415. doi: 10.1037/0022-0663.91.3.415

Wu, S., Packard, J. L., and Shu, H. (2009). "Morphological deficit and dyslexia subtypes in Chinese," in *Language Disorders in Speakers of Chinese*, eds S.-P. Law, B. S. Weekes, and A. M.-Y. Wong, (Bristol: Multilingual Matters), 112–137. doi: 10.21832/9781847691170-009

Xue, J., Shu, H., Li, H., Li, W., and Tian, X. (2013). The stability of literacy-related cognitive contributions to Chinese character naming and reading fluency. *J. Psycholinguist. Res.* 42, 433–450. doi: 10.1007/s10936-012-9228-0

Yeung, P., Ho, C., Chik, P. P., Lo, L., Luan, H., Chan, D., et al. (2011). Reading and spelling Chinese among beginning readers: what skills make a difference? *Sci. Stud. Read.* 15, 285–313. doi: 10.1080/10888438.2010.482149

Zhou, Y., Mcbride-Chang, C., Law, B. Y., Li, T., Cheung, C. Y., Wong, M. Y., et al. (2014). Development of reading-related skills in Chinese and English among Hong Kong Chinese children with and without dyslexia. *J. Exp. Child Psychol.* 122, 75–91. doi: 10.1016/j.jecp.2013.12.003

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., et al. (2010). Orthographic depth and its impact on universal predictors of reading: a cross-language investigation. *Psychol. Sci.* 21, 551–559. doi: 10.1177/0956797610363406

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Whole-Brain Functional Networks for Phonological and Orthographic Processing in Chinese Good and Poor Readers

Jing Yang[1] and Li Hai Tan[2,3]*

[1] *Bilingual Cognition and Development Lab, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China,* [2] *Center for Brain Disorders and Cognitive Science, Shenzhen University, Shenzhen, China,* [3] *Center for Language and Brain, Shenzhen Institute of Neuroscience, Shenzhen, China*

The neural basis of dyslexia in different languages remains unresolved, and it is unclear whether the phonological deficit as the core deficit of dyslexia is language-specific or universal. The current functional magnetic resonance imaging (fMRI) study using whole-brain data-driven network analyses investigated the neural mechanisms for phonological and orthographic processing in Chinese children with good and poor reading ability. Sixteen good readers and 16 poor readers were requested to make homophone judgments (phonological processing) and component judgments (visual-orthographic processing) of presented Chinese characters. Poor readers displayed worse performance than the good readers in phonological processing, but not in orthographic processing. Whole-brain activation analyses showed compensatory activations in the poor readers during phonological processing and automatic phonological production activation in the good readers during orthographic processing. Significant group differences in the topological properties of their brain networks were found only in orthographic processing. Analyses of nodal degree centrality and betweenness centrality revealed significant group differences in both phonological and orthographic processing. The present study supports the phonological core deficit hypothesis of reading difficulty in Chinese. It also suggests that Chinese good and poor readers might recruit different strategies and neural mechanisms for orthographic processing.

Keywords: dyslexia, phonological deficit, orthographic deficit, Chinese, functional brain network

INTRODUCTION

Developmental dyslexia, or in short, dyslexia, is characterized by a severe reading acquisition disorder that cannot be explained by general intelligence impairment, lack of education opportunities, or any sensory or neurological disorders (American Psychiatric Association, 2013). It is a widespread reading disorder that affects word recognition, decoding, and spelling abilities in 5–17% of the population, regardless of cultural or language backgrounds (Shaywitz et al., 1998; Ziegler et al., 2003; Siok et al., 2008; Cao et al., 2017). Phonological deficits, including impaired phonological representation and speech sound processing, are presented in the majority of dyslexics (Ziegler and Goswami, 2005) and therefore the phonological deficit hypothesis has been the most popular hypothesis about the cause of dyslexia (Rack et al., 1992; Pennington and Lefly, 2001; for a recent review, see Paulesu et al., 2014). This hypothesis posits that dyslexics are impaired

in their phonological representation and their ability to process and manipulate speech sounds (e.g., Shankweiler and Lundquist, 1992; Ziegler and Goswami, 2005), which adversely affects the development of mapping between written forms (graphemes) and speech sound (phonemes) and hinders reading development (Snowling, 1981, 1998; Muter et al., 1998; Ramus and Szenkovits, 2008; Hulme et al., 2012; Castles and Friedmann, 2014).

Phonological and Orthographic Deficits in Dyslexia

There is a tremendous amount of research on the brain mechanism of phonological processing deficits in dyslexics, and how such deficits affect reading development and might be relieved by phonological training (e.g., Shaywitz et al., 1998; Brunswick et al., 1999; Demb et al., 1999; Temple et al., 2001; Gaab et al., 2007; Cao et al., 2008; Tanaka et al., 2011; Steinbrink et al., 2012; Zhang et al., 2018). Most of the neuroimaging studies to date have investigated neural mechanism of dyslexia using visual word/pseudoword tasks and found reduced brain activation in the left temporo-parietal and temporo-occipital region in dyslexics speaking alphabetic languages (e.g., Rumsey et al., 1997; Paulesu et al., 2001, 2014; Schulz et al., 2009; van der Mark et al., 2009; Desroches et al., 2010; Pecini et al., 2011; Tanaka et al., 2011). The activation of the left inferior frontal gyrus (IFG) in dyslexics, however, increased in some studies (Shaywitz et al., 1998; Hoeft et al., 2007; MacSweeney et al., 2009) and decreased in other studies (Brambati et al., 2006; Cao et al., 2006; Booth et al., 2007; Wimmer and Schurz, 2010). Richlan et al. (2011) in a meta-analysis examined left temporo-parietal dysfunction for phonological deficits in dyslexic children and left ventral temporo-occipital dysfunction for visual-orthographic deficit in dyslexic adults. They found decreased activation of left ventral temporo-occipital region only in dyslexic adults.

Phonological deficits, however, are not the only problem in dyslexia. For example, Denckla and Rudel (1976) first reported picture naming problems in many people with dyslexia, who were slower than the normal when asked to rapidly name visual stimuli (for an overview, see Wolf et al., 2000). Wolf and Bowers, therefore, developed the double deficit hypothesis, which postulates that some people with dyslexia had a second independent naming speed deficit, which causes slower cross-modal matching of visual symbols and phonological codes, and therefore also causes reading problems (e.g., Bowers and Wolf, 1993; Wolf and Bowers, 1999; Vaessen et al., 2009).

Dyslexia is also suggested to be associated with orthographic deficits. First, rapid naming deficits seems to be quite universal among dyslexics in many languages, but phonological awareness deficit, difficulty to recognize and work with sounds in spoken language, are more common in opaque alphabetic languages (e.g., English) than transparent alphabetic languages (e.g., Italian) or non-alphabetic languages (e.g., Chinese) (e.g., Huang and Hanley, 1995; Ho et al., 2002; Ziegler et al., 2003; Tan et al., 2005; Vidyasagar and Pammer, 2010). Secondly, dyslexics exhibit deficits in processing letter strings in tasks with minimal phonological or lexical involvement, such as searching for a target letter in a string of consonants (e.g., Hawelka et al.,

2006; Bosse et al., 2007; Collis et al., 2013). Ziegler et al. (2010) reported that dyslexics performed significantly worse than age-matched controls with letter and digit strings but not with symbol strings. The authors suggest that these deficits cannot be explained by weak reading experience in dyslexics, or dysfunctional visual attention processing, and reflect a deficit in processing a string of letters in parallel, probably due to difficulty in the coding of letter position. Finally, some neuroimaging studies have also found dyslexics show less activation than the normals in left fusiform gyrus, a system specialized for processing the orthographic structure of well-learned visual word forms (Rumsey et al., 1997; Brunswick et al., 1999; Temple et al., 2001; Shaywitz and Shaywitz, 2003; Cao et al., 2006; van der Mark et al., 2009; Boros et al., 2016). For example, Desroches et al. (2010) reported reduced brain activation in the fusiform gyrus in dyslexics compared with the normals during an auditory rhyming task. The brain activation in left fusiform gyrus of the dyslexics correlated significantly and positively with their non-word reading performance. The authors, therefore, suggest that dyslexics were impaired in the access to orthography and the integration of orthographic and phonological processing.

The dysfunction activation of fusiform gyrus may be secondary to a primary dysfunction of the temporo-parietal region (Boros et al., 2016). Orthographic deficits in dyslexics increase the difficulty of selecting graphemes in fusiform gyrus, which are the input to the grapheme-phoneme processing and phonological decoding system in the temporo-parietal region. Therefore dyslexia might be characterized by the co-existence of orthographic and phonological processing difficulties (Siok et al., 2009).

Dyslexia in Chinese

Siok et al. (2004) found that Chinese dyslexic children reading in Chinese did not show underactivation in the left temporo-parietal regions as typically shown in studies of alphabetic languages. They reported reduced activity at Brodmann' area (BA) 9 in the left middle frontal gyrus (MFG), an area involved in syllabic processing of phonology (Siok et al., 2003). This study provides the first neural evidence to support previous findings of phonological awareness predicting reading development of Chinese children (e.g., McBride-Chang et al., 2008; Pan et al., 2011, 2015) and impaired phonological awareness in Chinese dyslexic children (e.g., Ho and Lai, 1999; Ziegler and Goswami, 2005), but also challenges the biological unity of dyslexia.

Unlike alphabetic languages, Chinese is a logographic language, in which the basic orthographic units, the characters, map onto morphemic meanings and to monosyllables with Chinese four tones in the spoken language. Therefore, Chinese reading needs a fine-grained visuospatial analysis to access characters' phonology and meaning. Chinese readers must learn the character phonology at the syllabic level as a whole by rote, and they might need additional strategies like writing to learn those characters (Siok et al., 2004; Tan et al., 2005; Ziegler, 2006; Cao et al., 2013; cf. Bi et al., 2009).

Siok et al. (2009) compared Chinese dyslexic children and normal children in a decision task of Chinese character physical size. The normal showed greater activation than the dyslexic

in the right inferior parietal lobe; the dyslexics, however, had more neural response than the normal participants in left inferior parietal lobe and lingual gyrus subserving visual analysis. According to the authors, phonological and orthographic disorders co-exist in the majority (83.33%) of Chinese dyslexics. The findings of Siok et al. (2009) are congruent with earlier behavioral reports of visual-orthographic deficits in Chinese dyslexics (Huang and Hanley, 1995; Ho et al., 2002).

Hu et al. (2010) examined brain activations of Chinese dyslexics, English dyslexics, English normal readers, and Chinese normal readers in a semantic decision task on written words. They found Chinese and English dyslexic adolescents had common underactivation than their normal controls in the left angular gyrus, left middle frontal, posterior temporal, and occipito-temporal regions. The authors suggest commonalities of manifestation of dyslexia in Chinese and English population, which could be influenced by readers' cognitive ability and learning environment, as is congruent with Ziegler's claim on the universal phonological core deficit of dyslexia (Ziegler, 2006).

Brain Connectivity in Dyslexia

A significant trend in cognitive neuroscience today is the brain connectivity approach, which explores the functional or structural connectivity patterns of brain regions that support cognitive or linguistic processing. A few studies have adopted this approach toward dyslexia.

In their pioneer work on dyslexia and connectivity, Horwitz et al. (1998), using positron emission tomography (PET) found that the dyslexics' left angular gyrus is functionally disconnected from the extrastriate occipital and temporal lobe regions during single-word reading, compared with the normal adults. They suggest a disconnected brain network in dyslexia. More recently, Boets et al. (2013) examined whether dyslexics' phonological deficits are caused by impaired phonological representation or by dysfunctional retrieval of phonological representations. They found that adult dyslexics have intact phonetic representations. Their functional and structural connectivity between the bilateral auditory cortices and the left IFG, however, is significantly smaller than the normal adults. Cao et al. (2017) focused on the phonological deficits of Chinese dyslexic children, who were asked to perform an auditory rhyming judgment task. They found that Chinese dyslexics were impaired in the left dorsal IFG and they had more reliance on the right precentral gyrus than the normal controls as a compensatory strategy. Their functional connectivity analyses showed that connectivity between the left STG and the left dorsal IFG was sensitive to task performance and/or reading skill rather than being dyslexic or not. In a functional connectivity study of orthographic processing of dyslexia, van der Mark et al. (2009) focused on the role of the left visual word form area in temporo-occipital area and found a significant disruption of the functional connectivity between the visual word face area (VWFA) and left inferior frontal and left inferior parietal language areas in the dyslexic children. They suggest that dyslexia is associated with impaired automatic visual word processing, along with deficits in orthographic and phonological processing. The studies mentioned above were based on the analysis of regions of interests (ROIs), and therefore

their results depend on the selected regions, which are arbitrary decisions by the authors. Finn et al. (2014) adopted a whole-brain, data-driven analysis to examine the functional networks in dyslexics. They found reduced connectivity in the visual word-form areas and increased right-hemisphere connectivity in the dyslexics compared with the normal adults. However, the parcellations in both the younger reader and older reader groups were generated from their groups of normal participants with limited group size (30–45). Their data analysis focused on group differences in regional connectivity and did not compare the topological features of brain networks.

The Present Study

The present functional magnetic resonance imaging (fMRI) study investigated phonological processing and orthographic processing in Chinese children with good and poor reading ability to improve the current understanding of the universal neural mechanism for dyslexia. All participants were asked to perform a homophone judgment task (phonological processing) and a component search task (visual-orthographic processing) inside the fMRI scanner. We examined group differences in their whole-brain activation and analyzed the topological features of their functional brain networks to reveal the neural mechanisms for phonological and visual-orthographic processing in Chinese good and poor readers.

MATERIALS AND METHODS

Participants

Five hundred and twenty-four 4th and 5th graders from the Beijing Yongtai Primary School in China participated in the screening for good readers and poor readers. Since there was no standardized dyslexia screening assessment or Chinese reading ability test in mainland China, we measured the children's reading ability using a character-reading test, their Chinese teachers' evaluation, and their school performance in the Chinese language course. This character-reading test was adapted from the reading test to evaluate Chinese children's reading ability by Tan et al. (2005), comprised 120 Chinese characters from the textbooks for third to fifth graders (40 characters for each grade) and 40 characters beyond the primary school textbooks. The 160 characters were printed on a standard A4 sheet, listed in 16 rows and 10 columns, and arranged from easy to difficult based on grade level. Each participant was asked to read out the 160 characters as accurately and as fast as possible with a time limit of 90 s. Their name accuracy (number of characters correctly named) represented their reading performance: Poor readers had reading scores 1.5 standard deviations below the mean; good readers had reading scores 1.5 standard deviations above the mean. Their reading performance was congruent with the evaluation from their Chinese teachers and their school performance in the Chinese course. Seventeen children with dyslexia and 16 controls participated in the present fMRI study. One participant from the normal group was excluded because of neurological disease found during the fMRI scans.

The reading performance (Mean $\pm$ SD = 115.75 $\pm$ 13.57) of the 16 participants in the normal group (9 men, average age = 10 years 1 month) was significantly better than that (Mean $\pm$ SD = 35.63 $\pm$ 13.59) of the 16 participants in the dyslexic group (12 men, average age = 10 years 6 months), t_{30} = 16.69, p < 0.001. All participants, who were native speakers of Chinese and right-handed (Oldfield, 1971), had average and matched non-verbal intelligence according to their performance in the Raven's Progressive Matrices (Raven et al., 1998) (good readers, Mean $\pm$ SD = 68.44 $\pm$ 15.78; poor readers, Mean $\pm$ SD = 75 $\pm$ 16.73; t_{30} = −1.141, p = 0.26). This fMRI study was approved by the Beijing Institutional Review Board at the Chinese Academy of Sciences. Written informed consent was obtained from each child and his/her legal guardians, mostly their parents.

Stimuli and Procedure

In this blocked-design fMRI study, both groups underwent a phonological session and a visual-orthographic session. During the phonological session, participants performed a homophone judgment task in experimental blocks: they were asked to judge whether the characters (e.g., "盐" sounds/yan2/and means "salt") presented had the same pronunciation including tones with the "pinyin[1]" (e.g., "yán" sounds/yan2/) specified at the instruction page before each experimental block. During the visual-orthographic session, participants completed a component judgment task: they were asked to identify whether the characters (e.g., "叔" sounds/shu1/and means "uncle") presented contained a radical (e.g., "又") specified at the instruction page before each experimental block. Chinese orthographic processing involves visuospatial analysis of Chinese characters and the application of orthographic rules (orthographic awareness). Component search task (orthographic search) asks participants to judge whether a character contained a designated a radical component and has been used as Chinese visual-orthographic processing task in previous studies (e.g., Siok and Fletcher, 2001; Ding et al., 2003).

Both sessions included four experimental blocks (homophone judgment/component search): each block began with a 2-s instruction and included eight trials; each trial started with a 500-ms presentation of Chinese character at the center of the screen, followed by a 2500-ms blank screen for responses. All the experimental blocks were interleaved with 12-s fixation blocks. Participants made "Yes" or "No" responses by clicking right or left buttons with their index fingers on a control box compatible with the fMRI scanner. The Chinese character stimuli, selected from the children's textbooks, were matched between experimental tasks in terms of character frequency and visual complexity (strokes).

MRI Acquisition

MRI images were acquired on a Siemens Vision Magnetom 3.0-tesla scanner with a circularly polarized head coil at the Beijing

---

[1]"Pinyin" is an alphabetic phonetic system mainly used in Mainland China to represent pronunciation of Chinese characters in Putonghua, standard spoken language in Mainland China. All children (6–7 years old) enrolled in primary school education at Mainland China are trained in Pinyin for 6–8 weeks before starting to learn Chinese characters.

MRI Imaging Center. Before the fMRI scans, all participants underwent a practice session and were visually familiarized with all the procedures and experimental conditions. They lay supine in the scanner with plastic ear-canal molds and looked up through a prism at a screen at the end of the scanner, while their heads were immobilized by a tightly fitting, vacuum pillow. A T_2*-weighted gradient-echo planar imaging (EPI) sequence was used for fMRI scans: slice thickness = 4 mm, in-plane resolution = 3.125 $\times$ 3.125 mm^2, and TR/TE/flip angle = 2000 ms/30 ms/90°. The field of view (FOV) was 200 $\times$ 200 mm^2, and the acquisition matrix was 64 $\times$ 64. Thirty-two contiguous axial slices were acquired parallel to the anterior commissure–posterior commissure (AC–PC) line covering the whole brain.

Data Analyses

Whole-Brain Activations

SPM 12 was used for image preprocessing and statistical analyses[2]. Functional images from each participant were realigned and normalized to an EPI template based on the ICBM152 stereotactic space, an approximation of canonical space (Talairach and Tournoux, 1988). The images were further re-sampled into 3 mm $\times$ 3 mm $\times$ 3 mm cubic voxels and spatially smoothed with an isotropic Gaussian kernel (6 mm full width at half-maximum). After motion-correction, the first three images (dummy images), corresponding to the period of transient hemodynamic change that occurred before the experimental trials, were discarded. The general linear model included 12 motion regressors was used to estimate the condition effect of each individual, while boxcar convolved with the canonical hemodynamic response function was selected as a reference function. Adjusted mean images were created for each condition after removing global signal and low-frequency covariates, using a high-pass filter with a cut-off of 128 s. Contrast images of homophone judgment minus fixation in phonological scanning session and component judgment minus fixation in visual-orthographic session were computed, using a Student's group t-test, which generated the statistical parametric maps of t-values. For each session, all the contrast estimates from dyslexic and normal groups were entered into a standard SPM second-level analysis with subjects treated as a random effect, using two-samples T-test to examine possible group differences in brain activations.

All the brain activations reported below were in MNI coordinate space and survived a corrected cluster-level threshold of p < 0.05 (single voxel p = 0.005, 10000 simulations, and a minimum cluster size of 25 voxels) using AlphaSim program in REST software (Song et al., 2011).

Network Construction

Functional brain networks for good readers and poor readers were constructed at the macroscale in which nodes represent brain regions, and edges present the statistical relationships of blood oxygenation level-dependent (BOLD) signals across different regions. Here, we used the 90 regions (45 for

---

[2]http://www.fil.ion.ucl.ac.uk/spm

each hemisphere) of the atlas of Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer et al., 2002) as nodes of the brain network. The averaged time series of all the voxels within each ROI was extracted in each individual. Edges, or interregional functional connectivity, were calculated using Pearson correlations between these regional task-related time series of all possible pairs of the 90 regions for each participant. The correlation coefficients were then transformed to z-scores via Fisher's transformation to improve normality (Lowe et al., 1998). Thus each participant has a 90×90 correlation matrix for phonological and visual-orthographic sessions, respectively.

Network Analysis

Threshold selection

We constructed binary undirected functional networks using a sparsity threshold ($5\% \leq$ sparsity $\leq 50\%$, interval = 5%) to comprehensively estimate topological properties covering a wide range of sparsity and remove spurious edges as much as possible (Yang et al., 2017; Zhu et al., 2017). Because the physiological interpretation of negative correlations is ambiguous (e.g., Murphy and Fox, 2017), functional connections with negative correlation values were not considered in the present analysis.

Network metrics

Our network analyses were performed in the GRETNA toolbox (Wang et al., 2015). We calculated both the global and node network metrics at each sparsity. These metrics included: (1) The "small-world" parameters of clustering coefficient (Cp), shortest path length (Lp), normalized clustering coefficient (γ), normalized shortest path length (λ), and small-worldness (σ); (2) Network efficiency measures of the local efficiency of the whole network (E_{loc}) and the global efficiency of the network (E_{glob}); (3) Nodal centrality degree and betweenness degree that reflect functional segregation and integration (Rubinov and Sporns, 2010).

Group comparisons based on topological metrics

To examine group differences of all the network metrics mentioned in the above section, two-sample t-test analyses were used for between-subject comparisons. To correct for multiple comparisons, we used a Bonferroni corrected threshold at the

significance level of 0.05. The network results were visualized using BrainNet Viewer (Xia et al., 2013).
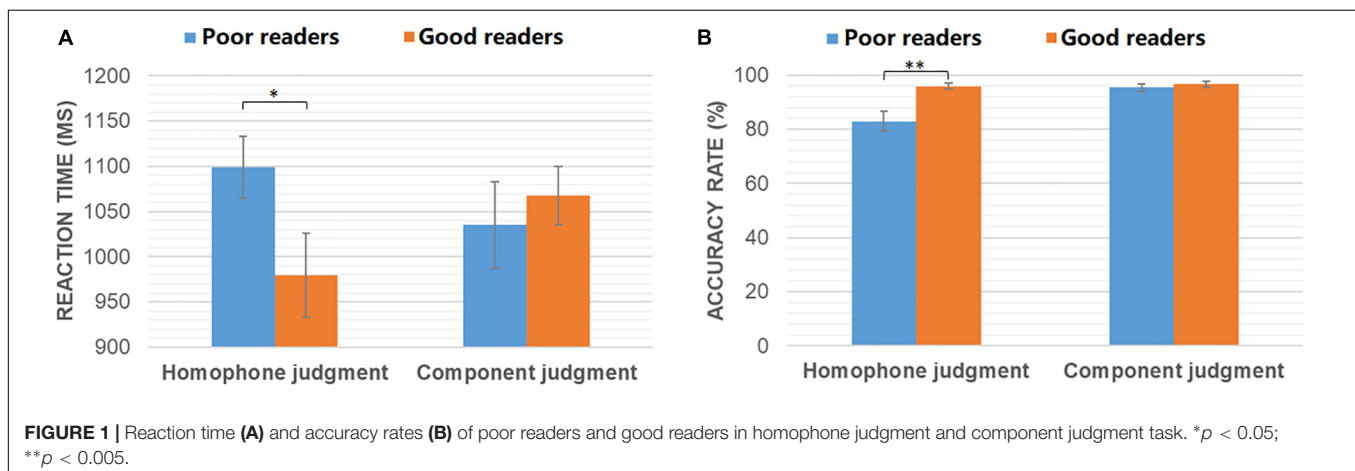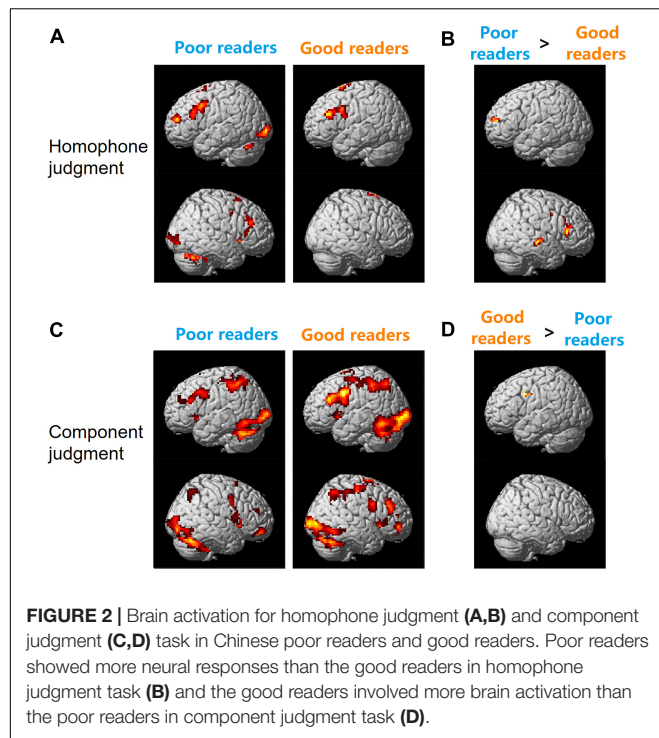
RESULTS

Behavioral Results

Independent-samples T-tests were conducted to compare the behavioral performance of good and poor readers in homophone judgment and component judgment tasks, respectively. As shown in **Figure 1**, poor readers were significantly slower ($t_{30} = -2.08$, $p = 0.046$) and less accurate ($t_{30} = 3.31$, $p = 0.004$) than the normals in the homophone judgment task. However, the two groups had similar performance in the component judgment task (ps > 0.05).

Whole-Brain Activations

As shown in **Figure 2A**, during the homophone judgment task, good readers recruited left MFG (BAs 9, 46), left IFG (pars triangularis, BA 45), and bilateral SMA (BAs 6, 8). In contrast, poor readers involved an extensive and symmetrical brain network, including the bilateral prefrontal cortex, insula, cingulate cortex, caudate nuclei, occipital regions, and cerebellum. Group comparisons showed poor readers had significantly more neural responses in the left anterior MFG, right IFG, right superior and middle temporal gyrus (MTG; **Figure 2B**). The good readers didn't show more neural responses compared with the poor readers.

During the component judgment task, Chinese good readers showed brain activations in bilateral middle and inferior frontal gyri, precentral gyri, SMA, insula, cingulate cortex, basal ganglia, and thalamus. Bilateral superior and inferior parietal lobules, posterior temporal-occipital cortex, and cerebellum were also involved in this group. The poor readers showed neural responses in those regions similar to that of the good readers (**Figure 2C**). During the component judgment task (in contrast to the fixation baseline condition), the good readers had significantly more neural activity in the left premotor cortex (BA 6) than the poor readers (**Figure 2D**). All reported group differences in brain activation were summarized in **Table 1**.

FIGURE 1 | Reaction time **(A)** and accuracy rates **(B)** of poor readers and good readers in homophone judgment and component judgment task. *$p < 0.05$; **$p < 0.005$.

FIGURE 2 | Brain activation for homophone judgment **(A,B)** and component judgment **(C,D)** task in Chinese poor readers and good readers. Poor readers showed more neural responses than the good readers in homophone judgment task **(B)** and the good readers involved more brain activation than the poor readers in component judgment task **(D)**.

Network Metrics

As shown in **Figures 3A–C**, significant group differences were found between their clustering coefficient (Cp), shortest path length (Lp), and normalized shortest path length (λ) of functional networks for visual-orthographic processing (component judgment task), but not for phonological processing (homophone judgment task). To be specific, during visual-orthographic processing, the brain networks of the dyslexic children displayed significantly higher Cp at the sparsity threshold of 45% (dyslexics, 0.76 ± 0.04; normal, 0.73 ± 0.02; $t = -2.14$, $p = 0.04$). They also had higher values of Lp than the normals for thresholds between 25 and 45%; the groups were significantly different in their λ at the thresholds of 30, 35, 40, and 45% (ps < 0.05).

Network Efficiency

For the homophone judgment task, there were no significant group differences in their local efficiency (E_{loc}) or global efficiency (E_{glob}). For the component judgment task, the good readers displayed higher global efficiency than the poor readers at the thresholds between 20 and 50% (ps < 0.05). No group difference was found for local efficiency in the component judgment task.

Nodal Centrality Degree

We used two-sample t-tests to examine group differences in nodal centrality measures of degree centrality and betweenness centrality at the strongest threshold (sparsity = 5%) so that all/most of the nodes were connected (**Table 2**). The poor readers displayed higher degree centrality in left middle temporal gyrus (MTG) during homophone judgment task compared with the good readers, who displayed higher degree centrality than the former in the right temporal pole (TP; superior and middle temporal gyri) during component judgment task (**Figure 4A**). As shown in **Figure 4B**, poor readers showed significantly higher betweenness centrality than the good readers in left calcarine fissure and right middle occipital gyrus in component judgment task. There were no significant group differences in betweenness centrality in homophone judgment task.
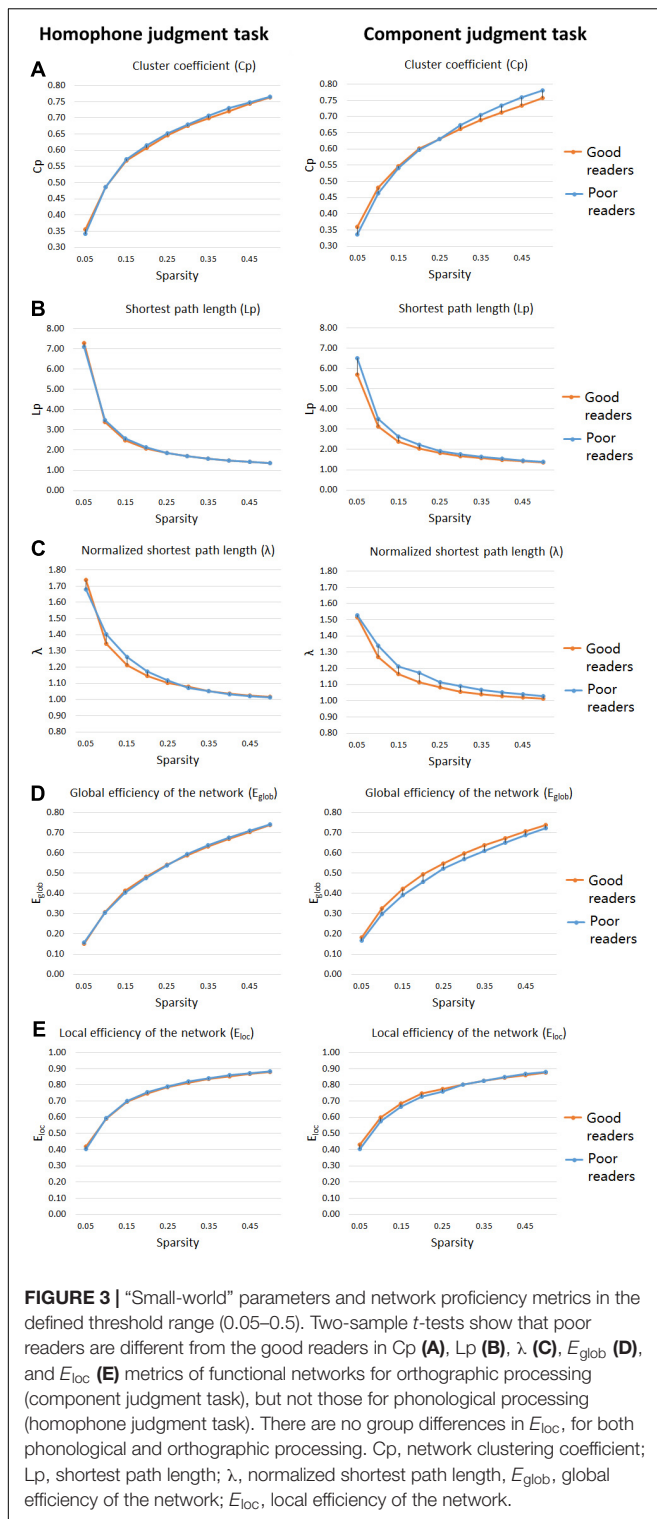
DISCUSSION

The present fMRI study using a whole-brain data-driven network approach examined the neural correlates of phonological and visual-orthographic processing in Chinese good readers and poor readers, who were forth or fifth graders matched in age and non-verbal intelligence. Our behavioral data showed that poor readers made more errors and responded more slowly than the good readers in phonological processing (homophone judgment task). There were no group differences in orthographic processing (component judgment task) at the behavioral level. Our behavioral findings are consistent with the phonological deficit hypothesis of dyslexia and suggest no orthographic deficits in Chinese children with reading difficulties (poor readers). Whole-brain activation analyses, however, revealed the poor readers compared with the good readers had hyperactivity in left MFG (BA 10), right IFG (BA 45), and right superior temporal sulcus (STS) (BA 22) during phonological processing, and hypoactivity in the left premotor cortex (BA 6) during visual-orthographic processing. In line with poor readers' behavioral deficits in phonological processing, the aberrant brain activations for phonological processing in Chinese poor readers suggests

TABLE 1 | Significant differences between Chinese poor readers and good readers in brain activations for homophone judgment and component judgment task.

Regions	L/R	BA	MNI			Z-value	L/R	BA	MNI			Z-value
			x	y	z				x	y	z	
Homophone judgment			Good readers > Poor readers						Poor readers > Good readers			
Middle frontal gyrus		None					L	10	−30	57	15	3.59
Inferior frontal gyrus							R	45	48	27	9	4.18
Superior temporal sulcus							R	22	66	−15	−3	3.5
Component judgment												
Premotor cortex	L	6	−48	6	36	3.19	None					

MNI, MNI coordinates. L, left hemisphere; R, right hemisphere.

FIGURE 3 | "Small-world" parameters and network proficiency metrics in the defined threshold range (0.05–0.5). Two-sample *t*-tests show that poor readers are different from the good readers in Cp **(A)**, Lp **(B)**, λ **(C)**, E_{glob} **(D)**, and E_{loc} **(E)** metrics of functional networks for orthographic processing (component judgment task), but not those for phonological processing (homophone judgment task). There are no group differences in E_{loc}, for both phonological and orthographic processing. Cp, network clustering coefficient; Lp, shortest path length; λ, normalized shortest path length, E_{glob}, global efficiency of the network; E_{loc}, local efficiency of the network.

the topological features of their functional brain networks. During phonological processing, there were no significant group differences in measures of functional segregation (cluster coefficient, Cp) or functional integration (shortest path length, Lp, or the normalized shortest path length, λ). Nor were they different in their values of the global efficiency (E_{glob}) or local efficiency (E_{loc}). In visual-orthographic processing, poor readers displayed larger functional segregation (Cp) and less functional integration (Lp, λ) than good readers, who showed higher global efficiency (E_{glob}).

Further analyses of node centrality showed that during phonological processing, poor readers had a larger value of degree centrality at the left posterior MTG (pMTG) than the good readers, implying its more interactive role as a hub in network of dyslexics. During visual-orthographic processing, the good readers showed more centrality degree in the right TP and less betweenness centrality in the left calcarine fissure and middle occipital gyrus. Based on previous findings and our data reported above, we suggest a phonological core deficit of Chinese dyslexia and different visual-orthographic processing mechanisms in Chinese good and poor readers.

Impaired Phonological Processing, More Efforts on Cognitive Control and Semantic Processing for an Intact Functional Brain Network

Consistent with previous reports on the phonological deficits of Chinese dyslexia, the present study showed that Chinese children with reading difficulties performed worse than the good readers in the homophone judgment task. We didn't find underactivation in either the left MFG, IFG, temporo-parietal region or fusiform gyrus in the poor readers as reported in previous studies of dyslexia, in particular, Chinese dyslexia (e.g., Shaywitz et al., 1998; Paulesu et al., 2001; Temple et al., 2001; Cao et al., 2008, 2017; Tanaka et al., 2011). Instead, hyperactivation was found in the left anterior prefrontal cortex (aPFC), right IFG, and right posterior STS (pSTS) in poor readers. The aPFC is responsible for integrating outcomes of separate cognitive operations in the pursuit of a higher behavioral goal (for a review, see Ramnani and Owen, 2004) and the right IFG is involved in cognitive control and is recruited when important cues are detected (e.g., Hampshire et al., 2010). Therefore, the larger involvement of left aPFC and right IFG might indicate Chinese poor readers recruited more cognitive control and outcome integration as a compensatory strategy, which is domain-general. Studies on dyslexia have reported reduced gray matter volume in dyslexic readers in the right STG and left STS (e.g., Richlan et al., 2013) and symmetrically distributed gray matter in STS (Dole et al., 2013).The underactivation of left temporo-parietal region is also well-documented in studies of dyslexia, especially in alphabetic languages (Rumsey et al., 1997; Paulesu et al., 2001, 2014; Schulz et al., 2009). We hypothesize that in addition to cognitive control and feedback strategies, our poor readers might have recruited the right homologous site of left pSTS for semantic association and memory, as the left pSTS is a cortical hub for semantic

neurological disorder underlying the phonological processing of dyslexics. For visual-orthographic processing, the two groups might both function normally with different neural correlates. To provide a complete picture of the brain connectivity profiles of Chinese children with reading difficulties, we examined

TABLE 2 | Significant differences between Chinese poor readers (Poor) and good readers (Good) in nodal degree centrality and betweenness centrality in homophone judgment task (Homophone) and component judgment task (Component) at the sparsity threshold of 5%. MNI, MNI coordinates.

Centrality measures	Node	Volume-based ROI (MNI)			Voxel size	T	P
		x (mm)	y (mm)	z (mm)			
Degree centrality							
Homophone: Poor > Good	Middle temporal gyrus	−56	−34	−2	1439	−2.28	0.03
Component: Good > Poor	Temporal pole: superior temporal gyrus	48	15	−17	400	2.33	0.026
	Temporal pole: middle temporal gyrus	44	15	−32	349	2.05	0.049
Betweenness centrality							
Component: Poor > Good	Calcarine fissure	−7	−79	6	648	−2.26	0.03
	Middle occipital gyrus	37	−80	19	595	−2.25	0.03

T, two-sample t-test; P, all p-values less than 0.05.

processing and the extraction of meaning from multiple sources of information (Liebenthal et al., 2014).

Although there were no significant group differences in the small-world properties (Cp, Lp, λ, E_{glob}, and E_{loc}) of their functional networks for phonological processing, the poor readers had a larger value of degree centrality than the good readers in the left pMTG, which contributes to controlled retrieval of conceptual knowledge (Davey et al., 2016). With this compensatory strategy, poor readers had similar global and local brain network efficiency, despite their poor performance in the phonological task.

Functioning Orthographic Processing, Less Automatic Phonological Retrieval and Multimodal Integration, More Delays in Visual Analysis Hub, and Low Efficient Brain Network

Our studies didn't find behavioral deficits of Chinese poor readers in visual-orthographic processing. However, they engaged different brain activation and functional network to complete the same task as good readers did. Specifically, when the poor readers were fully occupied by the visual-orthographic processing task, the good readers automatically and efficiently activated the phonology of the presented character stimulus, and displayed more brain activation in the left premotor cortex, which is involved in speech production, especially articulation (e.g., Small et al., 1998; Donnan et al., 1999).

It is possible that Chinese poor readers recruited different neural mechanisms for visual-orthographic processing because they tend to have larger values of cluster coefficient and shortest path length, which also brings them a disadvantage in global efficiency compared with the good readers. The global efficiency of a network is a measure of network integration (Achard and Bullmore, 2007; Rubinov and Sporns, 2010), implying poor readers have a lower integration of functional network for visual-orthographic processing. The degree centrality analysis showed within the functional brain network of poor readers, the centrality of the right TP is less than that of the good readers. As bilateral TP are the core neural substrate for the formation of semantic representation (e.g., Lambon Ralph et al., 2009), our studies seem to suggest that the semantic representation in poor readers are not informative or complete as in the good readers. Meanwhile, betweenness centrality analysis found bilateral posterior visual cortex (calcarine fissure and middle occipital gyrus) play a more active role in information transportation of poor readers than in that of good readers during visual-orthographic processing, which indeed suggests more dependence on the visual neural correlates when poor readers perform the same visual-orthographic task as the good readers.

Our findings of the abnormal functional network for orthographic processing in Chinese children with reading difficulties are consistent with previous findings on the topological organization of brain structural network in Chinese dyslexic children (Liu et al., 2015). The authors using a similar whole-brain network analysis approach examined the structural brain network of Chinese dyslexics and found higher local specialization, a tendency of lower E_{glob} and prolonged characteristic path length in the dyslexic than the normal,
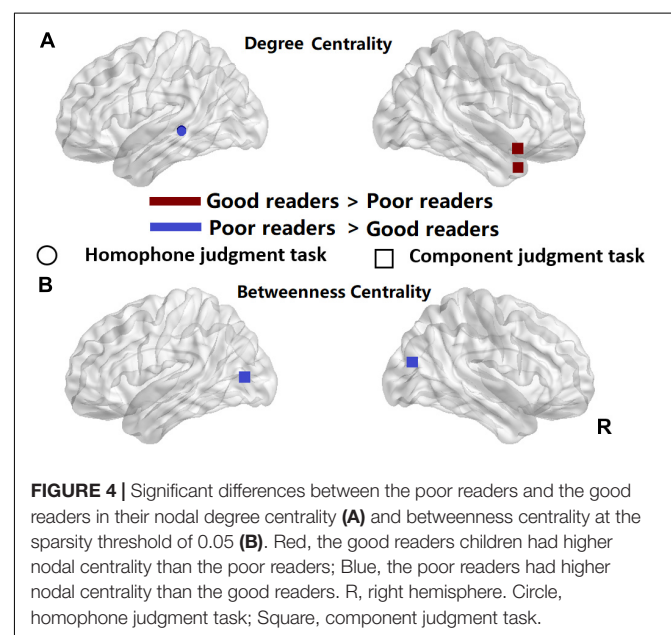

FIGURE 4 | Significant differences between the poor readers and the good readers in their nodal degree centrality **(A)** and betweenness centrality at the sparsity threshold of 0.05 **(B)**. Red, the good readers children had higher nodal centrality than the poor readers; Blue, the poor readers had higher nodal centrality than the good readers. R, right hemisphere. Circle, homophone judgment task; Square, component judgment task.

supporting our findings of the functional networks in Chinese children with reading difficulties.

Dynamic Brain Networks in Developmental Dyslexia

Using a whole-brain approach, the current study explored the differences between Chinese good and poor readers. Compared with previous studies in alphabetic languages, this study supports the phonological core deficit hypothesis of dyslexia and pointed out that behaviorally and neurologically dyslexics had manifestations of phonological processing deficits. Meanwhile, our results also imply distinct orthographic processing between Chinese good and poor readers, especially the inefficient functional brain network in poor readers during visual-orthographic processing.

The question remains: why abnormal brain activation and the inefficient brain functional network didn't cause orthographic processing deficits in Chinese dyslexics, as they do with phonological processing. We hypothesize that the neural mechanism for reading including the functional brain network is dynamic and developing, and behavioral performance of poor readers can be improved.

Training studies on dyslexia have provided numerous evidence on the effects of therapy or remediation on dyslexia. For example, the Tallal–Merzenich team provided intensive auditory training in dyslexic children and showed how the training rewired the children's brain (Merzenich et al., 1996; Tallal et al., 1996). Shaywitz et al. (2004) recruited second and third graders and administered phonologically mediated reading intervention to those with reading disabilities. Children who received the experimental intervention not only improved their reading performance but also showed increased brain activation in bilateral IFG, left STS, and temporo-occipital regions. Interestingly, Krafnick et al. (2011) reported gray matter volume changes in the left anterior fusiform gyrus/hippocampus, left precuneus, right hippocampus, and right anterior cerebellum during the intervention period. Those areas did not change after the training was stopped.

As we know, learning to read is associated with changes in brain activity. For example, Turkeltaub et al. (2003) in a cross-sectional fMRI study on subjects whose ages ranged from 6 to 22 years found reading acquisition is associated with increased activity in left MTG and IFG and decreased activity in the right inferior temporal regions. Learning to read also changes brain connectivity in dyslexics. Morken et al. (2017) traced reading process of dyslexics during their reading development. In this longitudinal study, participants were scanned through Pre literacy (6 years old), Emergent Literacy (8 years old), and Literacy (12 years old) stages. This study is the first fMRI study tracing the effectivity connectivity in dyslexics. Using Dynamic Causal Modelling (DCM) approach, they found different effectivity patterns in readers with and without dyslexia at age 6 and 8, but 12, implying by age 12, dyslexics reached functional, albeit poor reading skill with normalized effectivity close to the normal.

In the current study, participants were fourth and fifth graders, who had at least 5 years of experience in Chinese character writing and their Chinese literacy is close to the Literacy stage in Morken et al. (2017). It is possible that poor readers have orthographic deficits in their early years of Chinese reading acquisition. After they begin to receive school education, they are asked to do a lot of practice on Chinese writing and spelling to memorize Chinese words by rote in school and after school. Not surprisingly, Chinese writing can predict children's reading development (e.g., Tan et al., 2005; Cao et al., 2013). With reading development and intensive writing practice, their visual-orthographic processing, which was at a disadvantage in the beginning, might be improved to the extent that the differences between good and poor readers are not significant in terms of their behavioral performance. Only by neuroimaging techniques, we were able to reveal group differences in their neural substrates for visual-orthographic processing.

Meanwhile, the phonological deficit as the core deficit of dyslexia is not alleviated as reading skill approve. Their behavioral performance in phonological manipulation is still significantly different from good readers. Most of the intervention studies on dyslexia adopt the phonological-based training program. If more phonological-based training is used in the classroom setting, phonological deficits might be less in dyslexics as their reading literacy increases.

CONCLUSION

This study used a whole-brain data-driven network approach to examine the topological features of functional brain networks for phonological and visual-orthographic processing in Chinese good and poor readers. Our results suggest phonological deficits and aberrant neural mechanisms in Chinese poor readers, implying a language-universal phonological deficit in dyslexia. Our findings also indicate good and poor readers rely on different neural mechanisms or strategies in visual-orthographic processing to arrive at similar behavioral performance. To fully understand how phonological processing and visual-orthographic processing progress as reading literacy develops, we will need longitudinal studies tracking the reading development of dyslexics in typical classroom settings using brain imaging techniques.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Key Laboratory of Brain and Cognitive Sciences at The University of Hong Kong. Written informed consent to

participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

LT and JY conceived the presented study, collected the data, discussed the results, and contributed to the final manuscript. JY performed the data analyses.

REFERENCES

Achard, S., and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comp. Biol.* 3:e17. doi: 10.1371/journal.pcbi.0030017

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th edn).* Arlington, VA: American Psychiatric Association.

Bi, Y., Han, Z., and Zhang, Y. (2009). Reading does not depend on writing, even in Chinese. *Neuropsychologia* 47, 1193–1199. doi: 10.1016/j.neuropsychologia.2008.11.006

Boets, B., de Beeck, H. P. O., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342, 1251–1254. doi: 10.1126/science.1244333

Booth, J. R., Wood, L., Lu, D., Houk, J. C., and Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain Res.* 1133, 136–144. doi: 10.1016/j.brainres.2006.11.074

Boros, M., Anton, J.-L., Pech-Georgel, C., Grainger, J., Szwed, M., and Ziegler, J. C. (2016). Orthographic processing deficits in developmental dyslexia: beyond the ventral visual stream. *NeuroImage* 128, 316–327. doi: 10.1016/j.neuroimage.2016.01.014

Bosse, M.-L., Tainturier, M. J., and Valdois, S. (2007). Developmental dyslexia: the visual attention span deficit hypothesis. *Cognition* 104, 198–230. doi: 10.1016/j.cognition.2006.05.009

Bowers, P. G., and Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Read. Writ.* 5, 69–85. doi: 10.1007/bf01026919

Brambati, S. M., Termine, C., Ruffino, M., Danna, M., Lanzi, G., Stella, G., et al. (2006). Neuropsychological deficits and neural dysfunction in familial dyslexia. *Brain Res.* 1113, 174–185. doi: 10.1016/j.brainres.2006.06.099

Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., and Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: a search for Wernicke's Wortschatz? *Brain* 122, 1901–1917. doi: 10.1093/brain/122.10.1901

Cao, F., Bitan, T., and Booth, J. R. (2008). Effective brain connectivity in children with reading difficulties during phonological processing. *Brain Lang.* 107, 91–101. doi: 10.1016/j.bandl.2007.12.009

Cao, F., Bitan, T., Chou, T.-L., Burman, D. D., and Booth, J. R. (2006). Deficient orthographic and phonological representations in children with dyslexia revealed by brain activation patterns. *J. Child Psychol. Psychiatry* 47, 1041–1050. doi: 10.1111/j.1469-7610.2006.01684.x

Cao, F., Vu, M., Lung Chan, D. H., Lawrence, J. M., Harris, L. N., Guan, Q., et al. (2013). Writing affects the brain network of reading in Chinese: a functional magnetic resonance imaging study. *Hum. Brain Mapp.* 34, 1670–1684. doi: 10.1002/hbm.22017

Cao, F., Yan, X., Wang, Z., Liu, Y., Wang, J., Spray, G. J., et al. (2017). Neural signatures of phonological deficits in Chinese developmental dyslexia. *Neuroimage* 146, 301–311. doi: 10.1016/j.neuroimage.2016.11.051

Castles, A., and Friedmann, N. (2014). Developmental dyslexia and the phonological deficit hypothesis. *Mind Lang.* 29, 270–285. doi: 10.1111/mila.12050

Collis, N. L., Kohnen, S., and Kinoshita, S. (2013). The role of visual spatial attention in adult developmental dyslexia. *Quar. J. Exp. Psychol.* 66, 245–260. doi: 10.1080/17470218.2012.705305

Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., et al. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: integration of anterior temporal lobe with executive processes. *Neuroimage* 137, 165–177. doi: 10.1016/j.neuroimage.2016.05.051

Demb, J. B., Poldrack, R. A., and Gabrieli, J. D. (1999). "Functional neuroimaging of word processing in normal and dyslexic readers," in *Language, speech, and communication. Converging Methods for Understanding Reading and Dyslexia*, eds R. M. Klein, and P. A. McMullen (Cambridge, MA: The MIT Press), 245–304.

Denckla, M. B., and Rudel, R. G. (1976). Naming of object-drawings by dyslexic and other learning disabled children. *Brain Lang.* 3, 1–15. doi: 10.1016/0093-934X(76)90001-8

Desroches, A. S., Cone, N. E., Bolger, D. J., Bitan, T., Burman, D. D., and Booth, J. R. (2010). Children with reading difficulties show differences in brain regions associated with orthographic processing during spoken language processing. *Brain Res.* 1356, 73–84. doi: 10.1016/j.brainres.2010.07.097

Ding, G., Perry, C., Peng, D., Ma, L., Li, D., Xu, S., et al. (2003). Neural mechanisms underlying semantic and orthographic processing in Chinese–English bilinguals. *NeuroReport* 14, 1557–1562. doi: 10.1097/00001756-200308260-00003

Dole, M., Meunier, F., and Hoen, M. (2013). Gray and white matter distribution in dyslexia: a VBM study of superior temporal gyrus asymmetry. *PLoS One* 8:e76823. doi: 10.1371/journal.pone.0076823

Donnan, G. A., Carey, L. M., and Saling, M. M. (1999). More (or less) on Broca. *Lancet* 353, 1031–1032. doi: 10.1016/s0140-6736(99)90052-1

Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., et al. (2014). Disruption of functional networks in dyslexia: a whole-brain, data-driven analysis of connectivity. *Biol. Psychiatry* 76, 397–404. doi: 10.1016/j.biopsych.2013.08.031

Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., and Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: an fMRI study. *Restorative Neurol. Neurosci.* 25, 295–310.

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., and Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage* 50, 1313–1319. doi: 10.1016/j.neuroimage.2009.12.109

Hawelka, S., Huber, C., and Wimmer, H. (2006). Impaired visual processing of letter and digit strings in adult dyslexic readers. *Vis. Res.* 46, 718–723. doi: 10.1016/j.visres.2005.09.017

Ho, C. S.-H., Chan, D. W.-O., Tsang, S.-M., and Lee, S.-H. (2002). The cognitive profile and multiple-deficit hypothesis in Chinese developmental dyslexia. *Dev. Psychol.* 38:543. doi: 10.1037//0012-1649.38.4.543

Ho, C. S.-H., and Lai, D. N.-C. (1999). Naming-speed deficits and phonological memory deficits in Chinese developmental dyslexia. *Learn. Individ. Differ.* 11, 173–186. doi: 10.1016/s1041-6080(00)80004-7

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., et al. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4234–4239. doi: 10.1073/pnas.0609399104

Horwitz, B., Rumsey, J. M., and Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 95, 8939–8944. doi: 10.1073/pnas.95.15.8939

Hu, W., Lee, H. L., Zhang, Q., Liu, T., Geng, L. B., Seghier, M. L., et al. (2010). Developmental dyslexia in Chinese and English populations: dissociating the effect of dyslexia from language differences. *Brain* 133, 1694–1706. doi: 10.1093/brain/awq106

FUNDING

Huang, H.-S., and Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition* 54, 73–98. doi: 10.1016/0010-0277(94)00641-w

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., and Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: combining intervention studies with mediation analyses. *Psychol. Sci.* 23, 572–577. doi: 10.1177/0956797611435921

Krafnick, A. J., Flowers, D. L., Napoliello, E. M., and Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *Neuroimage* 57, 733–741. doi: 10.1016/j.neuroimage.2010.10.062

Lambon Ralph, M. A., Pobric, G., and Jefferies, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: convergent evidence from rTMS. *Cereb. Cortex* 19, 832–838. doi: 10.1093/cercor/bhn131

Liebenthal, E., Desai, R. H., Humphries, C., Sabri, M., and Desai, A. (2014). The functional organization of the left STS: a large scale meta-analysis of PET and fMRI studies of healthy adults. *Front. Neurosci.* 8:289. doi: 10.3389/fnins.2014.00289

Liu, K., Shi, L., Chen, F., Waye, M. M., Lim, C. K., Cheng, P. W., et al. (2015). Altered topological organization of brain structural network in Chinese children with developmental dyslexia. *Neurosci. Lett.* 589, 169–175. doi: 10.1016/j.neulet.2015.01.037

Lowe, M. J., Mock, B. J., and Sorenson, J. A. (1998). Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. *Neuroimage* 7, 119–132. doi: 10.1006/nimg.1997.0315

MacSweeney, M., Brammer, M. J., Waters, D., and Goswami, U. (2009). Enhanced activation of the left inferior frontal gyrus in deaf and dyslexic adults during rhyming. *Brain* 132, 1928–1940. doi: 10.1093/brain/awp129

McBride-Chang, C., Tardif, T., Cho, J.-R., Shu, H. U. A., Fletcher, P., Stokes, S. F., et al. (2008). What's in a word? morphological awareness and vocabulary knowledge in three languages. *Appl. Psycholinguist.* 29, 437–462. doi: 10.1017/s014271640808020x

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., and Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* 271, 77–81. doi: 10.1126/science.271.5245.77

Morken, F., Helland, T., Hugdahl, K., and Specht, K. (2017). Reading in dyslexia across literacy development: a longitudinal study of effective connectivity. *Neuroimage* 144, 92–100. doi: 10.1016/j.neuroimage.2016.09.060

Murphy, K., and Fox, M. D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154, 169–173. doi: 10.1016/j.neuroimage.2016.11.052

Muter, V., Hulme, C., Snowling, M., and Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *J. Exp. Child Psychol.* 71, 3–27. doi: 10.1006/jecp.1998.2453

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Pan, J., McBride-Chang, C., Shu, H., Liu, H., Zhang, Y., and Li, H. (2011). What is in the naming? a 5-year longitudinal study of early rapid naming and phonological sensitivity in relation to subsequent reading skills in both native Chinese and English as a second language. *J. Educ. Psychol.* 103:897. doi: 10.1037/a0024344

Pan, J., Shu, H., Wang, Y., and Yan, M. (2015). Parafoveal activation of sign translation previews among deaf readers during the reading of Chinese sentences. *Mem. Cogn.* 43, 964–972. doi: 10.3758/s13421-015-0511-9

Paulesu, E., Danelli, L., and Berlingeri, M. (2014). Reading the dyslexic brain: multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Front. Hum. Neurosci.* 8:830. doi: 10.3389/fnhum.2014.00830

Paulesu, E., Démonet, J.-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: cultural diversity and biological unity. *Science* 291, 2165–2167. doi: 10.1126/science.1057179

Pecini, C., Biagi, L., Brizzolara, D., Cipriani, P., Di Lieto, M. C., Guzzetta, A., et al. (2011). How many functional brains in developmental dyslexia? When the history of language delay makes the difference. *Cogn. Behav. Neurol.* 24, 85–92. doi: 10.1097/WNN.0b013e318222a4c2

Pennington, B. F., and Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Dev.* 72, 816–833. doi: 10.1111/1467-8624.00317

Rack, J. P., Snowling, M. J., and Olson, R. K. (1992). developmental dyslexia: a review. *Read. Res. Quar.* 27, 29–53.

Ramnani, N., and Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat. Rev. Neurosci.* 5:184. doi: 10.1038/nrn1343

Ramus, F., and Szenkovits, G. (2008). What phonological deficit? *Quar. J. Exp. Psychol.* 61, 129–141. doi: 10.1080/17470210701508822

Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven'sProgressive Matrices and Vocabulary Scales.* San Antonio, TX: Pearson,Inc.

Richlan, F., Kronbichler, M., and Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage* 56, 1735–1742. doi: 10.1016/j.neuroimage.2011.02.040

Richlan, F., Kronbichler, M., and Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies. *Hum. Brain Mapp.* 34, 3055–3065. doi: 10.1002/hbm.22127

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Rumsey, J. M., Nace, K., Donohue, B., Wise, D., Maisog, J. M., and Andreason, P. (1997). A positron emission tomographic study of impaired word recognition and phonological processing in dyslexic men. *Arch. Neurol.* 54, 562–573. doi: 10.1001/archneur.1997.00550170042013

Schulz, E., Maurer, U., van der Mark, S., Bucher, K., Brem, S., Martin, E., et al. (2009). Reading for meaning in dyslexic and young children: distinct neural pathways but common endpoints. *Neuropsychologia* 47, 2544–2557. doi: 10.1016/j.neuropsychologia.2009.04.028

Shankweiler, D., and Lundquist, E. (1992). On the relations between learning to spell and learning to read. *Adv. Psychol.* 94, 179–192. doi: 10.1016/s0166-4115(08)62795-8

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biol. Psychiatry* 55, 926–933. doi: 10.1016/j.biopsych.2003.12.019

Shaywitz, S. E., and Shaywitz, B. A. (2003). Dyslexia (specific reading disability). *Pediatr. Rev.* 24, 147–153. doi: 10.1542/pir.24-5-147

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., et al. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 95, 2636–2641.

Siok, W. T., and Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Dev. Psychol.* 37:886. doi: 10.1037/0012-1649.37.6.886

Siok, W. T., Jin, Z., Fletcher, P., and Tan, L. H. (2003). Distinct brain regions associated with syllable and phoneme. *Hum. Brain Mapp.* 18, 201–207. doi: 10.1002/hbm.10094

Siok, W. T., Niu, Z., Jin, Z., Perfetti, C. A., and Tan, L. H. (2008). A structural–functional basis for dyslexia in the cortex of Chinese readers. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5561–5566. doi: 10.1073/pnas.0801750105

Siok, W. T., Perfetti, C. A., Jin, Z., and Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature* 431:71. doi: 10.1038/nature02865

Siok, W. T., Spinks, J. A., Jin, Z., and Tan, L. H. (2009). Developmental dyslexia is characterized by the co-existence of visuospatial and phonological disorders in Chinese children. *Curr. Biol.* 19, R890–R892. doi: 10.1016/j.cub.2009.08.014

Small, S. L., Flores, D. K., and Noll, D. C. (1998). Different neural circuits subserve reading before and after therapy for acquired dyslexia. *Brain Lang.* 62, 298–308. doi: 10.1006/brln.1998.1951

Snowling, M. (1998). Dyslexia as a phonological deficit: evidence and implications. *Child Psychol. Psychiatr. Rev.* 3, 4–11. doi: 10.1111/1475-3588.00201

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychol. Res.* 43, 219–234. doi: 10.1007/bf00309831

Song, X.-W., Dong, Z.-Y., Long, X.-Y., Li, S.-F., Zuo, X.-N., Zhu, C.-Z., et al. (2011). REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS One* 6:e25031. doi: 10.1371/journal.pone.0025031

Steinbrink, C., Groth, K., Lachmann, T., and Riecker, A. (2012). Neural correlates of temporal auditory processing in developmental dyslexia during German vowel length discrimination: an fMRI study. *Brain Lang.* 121, 1–11. doi: 10.1016/j.bandl.2011.12.003

Talairach, J., and Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain.* New York, NY: Thieme.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271, 81–84. doi: 10.1126/science.271.5245.81

Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., and Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8781–8785.

Tanaka, H., Black, J. M., Hulme, C., Stanley, L. M., Kesler, S. R., Whitfield-Gabrieli, S., et al. (2011). The brain basis of the phonological deficit in dyslexia is independent of IQ. *Psychol. Sci.* 22, 1442–1451. doi: 10.1177/0956797611419521

Temple, E., Poldrack, R. A., Salidis, J., Deutsch, G. K., Tallal, P., Merzenich, M. M., et al. (2001). Disrupted neural responses to phonological and orthographic processing in dyslexic children: an fMRI study. *Neuroreport* 12, 299–307. doi: 10.1097/00001756-200102120-00024

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., and Eden, G. F. (2003). Development of neural mechanisms for reading. *Nat. Neurosci.* 6:767.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Vaessen, A., Gerretsen, P., and Blomert, L. (2009). Naming problems do not reflect a second independent core deficit in dyslexia: double deficits explored. *J. Exp. Child Psychol.* 103, 202–221. doi: 10.1016/j.jecp.2008.12.004

van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmüller, J., et al. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *Neuroimage* 47, 1940–1949. doi: 10.1016/j.neuroimage.2009.05.021

Vidyasagar, T. R., and Pammer, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends Cogn. Sci.* 14, 57–63. doi: 10.1016/j.tics.2009.12.003

Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., and He, Y. (2015). GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. *Front. Hum. Neurosci.* 9:386. doi: 10.3389/fnhum.2015.00386

Wimmer, H., and Schurz, M. (2010). Dyslexia in regular orthographies: manifestation and causation. *Dyslexia* 16, 283–299. doi: 10.1002/dys.411

Wolf, M., and Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *J. Educ. Psychol.* 91:415. doi: 10.1037//0022-0663.91.3.415

Wolf, M., Bowers, P. G., and Biddle, K. (2000). Naming-speed processes, timing, and reading: a conceptual review. *J. Learn. Disabil.* 33, 387–407. doi: 10.1177/002221940003300409

Xia, M., Wang, J., and He, Y. (2013). BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910

Yang, X., Liu, J., Meng, Y., Xia, M., Cui, Z., Wu, X., et al. (2017). Network analysis reveals disrupted functional brain circuitry in drug-naive social anxiety disorder. *Neuroimage* 190, 213–223. doi: 10.1016/j.neuroimage.2017.12.011

Zhang, M., Xie, W., Xu, Y., and Meng, X. (2018). Auditory temporal perceptual learning and transfer in Chinese-speaking children with developmental dyslexia. *Res. Dev. Disabil.* 74, 146–159. doi: 10.1016/j.ridd.2018.01.005

Zhu, H., Qiu, C., Meng, Y., Yuan, M., Zhang, Y., Ren, Z., et al. (2017). Altered topological properties of brain networks in social anxiety disorder: a resting-state functional MRI study. *Sci. Rep.* 7:43089. doi: 10.1038/srep43089

Ziegler, J. C. (2006). Do differences in brain activation challenge universal theories of dyslexia? *Brain Lang.* 98, 341–343. doi: 10.1016/j.bandl.2005.05.002

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131:3. doi: 10.1037/0033-2909.131.1.3

Ziegler, J. C., Pech-Georgel, C., Dufau, S., and Grainger, J. (2010). Rapid processing of letters, digits and symbols: what purely visual-attentional deficit in developmental dyslexia? *Dev. Sci.* 13, F8–F14. doi: 10.1111/j.1467-7687.2010.00983.x

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., and Schulte-Körne, G. (2003). Developmental dyslexia in different languages: language-specific or universal? *J. Exp. Child Psychol.* 86, 169–193. doi: 10.1016/s0022-0965(03)00139-5

Check for
updates

The Functional Neuroanatomy of Developmental Dyslexia Across Languages and Writing Systems

Fabio Richlan*

Centre for Cognitive Neuroscience, Department of Psychology, Paris Lodron University of Salzburg, Salzburg, Austria

The present article reviews the literature on the functional neuroanatomy of developmental dyslexia across languages and writing systems. This includes comparisons of alphabetic languages differing in orthographic depth as well as comparisons across alphabetic, syllabic, and logographic writing systems. It provides a synthesis of the evidence for both universal and language-specific effects on dyslexic functional brain activation abnormalities during reading and reading-related tasks. Specifically, universal reading-related underactivation of dyslexic readers relative to typical readers is identified in core regions of the left hemisphere reading network including the occipito-temporal, temporo-parietal, and inferior frontal cortex. Orthography-specific dyslexic brain abnormalities are mainly related to the degree and spatial extent of under- and overactivation clusters. In addition, dyslexic structural gray matter abnormalities across languages and writing systems are analyzed. The neuroimaging findings are linked to the universal and orthography-dependent behavioral manifestations of developmental dyslexia. Finally, the present article provides insights into potential compensatory mechanisms that may support remediation across languages and writing systems.

Keywords: brain, development, dyslexia, language, magnetic resonance imaging, orthography, reading, writing

INTRODUCTION

Developmental dyslexia is a disorder characterized by severe and persistent problems in the acquisition of literacy. Performance in reading skills is markedly below the age-adequate norm – in the absence of problems regarding intelligence, motivation, vision, or educational environment (American Psychiatric Association, 2013; World Health Organization, 2016). It has become evident from numerous studies that developmental dyslexia may not be viewed as a simple, single-trait disorder, that is, no single behavioral phenotype can be considered as a "typical" manifestation of dyslexia. There are problems in diverse aspects of literacy including reading fluency, accuracy, comprehension, and/or spelling, and people affected by dyslexia often present a mixture of different severities of these problems (e.g., Lyon et al., 2003). In addition, problems in learning to read are often comorbid with atypical or delayed oral language development (e.g., Catts et al., 2008; Peterson et al., 2009), writing disabilities, attention-deficit hyperactivity disorder (ADHD), and math disabilities/dyscalculia (e.g., Landerl and Moll, 2010; Willcutt et al., 2010).

Using neuroimaging techniques such as functional magnetic resonance imaging (fMRI), cognitive neuroscientific research has identified brain circuits crucially involved in typical and dyslexic reading. These studies have converged on a coarse functional neuroanatomical model of reading and developmental dyslexia. The model proposes abnormal brain activation in dyslexic readers in the left posterior temporo-parietal (TP) cortex (middle temporal gyrus, superior temporal gyrus, supramarginal gyrus, and angular gyrus), the left occipito-temporal (OT) cortex (inferior temporal gyrus and fusiform gyrus), and the left frontal cortex (inferior frontal gyrus and precentral gyrus).

As identified by meta-analyses, the most consistent finding is dyslexic underactivation relative to typical readers in the left TP and OT cortex. In addition, dyslexic underactivation was identified in the left inferior frontal gyrus (IFG) and dyslexic overactivation in the adjacent left precentral gyrus (Richlan et al., 2009, 2011; Martin et al., 2016; Hancock et al., 2017). Occasional reports on other bilateral cortical, subcortical, and cerebellar dyslexic deficits are not supported by the meta-analyses. Obviously, these dyslexic activation abnormalities depend largely on the utilized functional activation tasks during brain scanning, which are often targeted at providing evidence in favor of a specific neurocognitive deficit theory of dyslexia.

Although we have convincing evidence that the functioning of the above mentioned left TP, OT, and IFG cortical regions is altered in developmental dyslexia during reading and reading-related tasks, it is still an open question how the presumed functional and gray matter (GM) structural impairments in these regions lead to the severe and persistent reading problems of dyslexic readers. In other words, the question is not so much of whether and if so, where in the brain dyslexic abnormalities exist, but rather on how these brain regions might underlie reading- and spelling-related cognitive processes in typical and dyslexic readers. The present review article aims at providing an integrative overview and synopsis of the functional and structural brain abnormalities in dyslexic readers across languages and writing systems.

Specifically, the goal here is to focus on functional activation and GM structure; and on the commonalities and differences in these measures in developmental dyslexia across languages and writing systems. First, the functional neuroanatomy of developmental dyslexia across alphabetic languages differing in orthographic depth will be discussed. Second, the neurobiology of developmental dyslexia will be compared across alphabetic, syllabic, and logographic writing systems. Third, GM structural brain abnormalities in developmental dyslexia will be discussed. Finally, there will be a section on potential compensatory mechanisms that may support remediation across languages and writing systems.

Research on the relationship between functional activation and GM structure and their effects on reading development is of crucial importance but still scarce. Therefore, innovative approaches using intervention studies and longitudinal research will also be discussed. With respect to functional and structural connectivity in developmental dyslexia – which is beyond the scope of the present review – the reader is referred to other recent studies and meta-analyses (e.g., Ben-Shachar et al., 2007;

Cao et al., 2008, 2017; van der Mark et al., 2011; Vandermosten et al., 2012; Koyama et al., 2013; Dehaene et al., 2015; Olulade et al., 2015; Schurz et al., 2015; Alvarez and Fiez, 2018).

THE FUNCTIONAL NEUROANATOMY OF DEVELOPMENTAL DYSLEXIA ACROSS ALPHABETIC LANGUAGES DIFFERING IN ORTHOGRAPHIC DEPTH

Orthographic depth (OD) (i.e., the complexity, consistency, or transparency of grapheme-phoneme correspondences in written alphabetic language) (Frost et al., 1987) is a well-known factor influencing the acquisition of fast and accurate reading (Seymour et al., 2003; Landerl et al., 2013). Correspondingly, the behavioral manifestations of developmental dyslexia vary as a function of OD. Specifically, inaccurate mapping from graphemes to the corresponding phonemes is a particular hallmark of developmental dyslexia in irregular or deep orthographies – especially for English. On the contrary, persistent slow and dysfluent word recognition is a universal characteristic of developmental dyslexia across all alphabetic orthographies. Here we examine the question of how the different behavioral manifestations of developmental dyslexia are reflected in the functional neuroanatomical patterns identified by brain imaging studies.

The predominant view proposed a "cultural diversity and biological unity" account of developmental dyslexia, claiming a universal neurocognitive basis of the disorder across languages. This position was based on a seminal PET study comparing the brain activation of Italian, French, and English adult dyslexic readers in response to explicit and implicit reading tasks (Paulesu et al., 2001). The universal neurobiological substrate of developmental dyslexia across languages was reflected in underactivation (relative to typical readers) in a large left hemisphere cluster comprising the superior temporal gyrus (STG), middle temporal gyrus (MTG), inferior temporal gyrus (ITG), and middle occipital gyrus (MOG). Crucially, no orthography-specific effects in reading-related brain activation were identified in the direct statistical comparison of the dyslexic readers from the three languages varying in OD.

A qualitative summary and critical discussion of the Paulesu et al. (2001) study and more recent cross-linguistic brain imaging studies provided additional orthography-specific predictions regarding the degree and spatial extent of dyslexic under- and overactivation clusters relative to typical readers (Richlan, 2014). Together with the universal dysfunctions in core regions of the left hemisphere reading network (Pugh et al., 2005; Richlan, 2012; Martin et al., 2015), the presumed orthography-specific effects were derived from different functional neuroanatomical models of developmental dyslexia and dependent on the particular characteristics and processing demands of the language. In addition to differences in regional brain activation, deep orthographies (DO) and shallow orthographies (SO) were proposed to be associated with differences in the functional and effective connectivity between brain regions (Schurz et al., 2015).

Consequently, Martin et al. (2016) used coordinate-based meta-analysis in order to investigate the universal and orthography-specific predictions regarding dyslexic brain activation. Specifically, commonalities and differences of dyslexic functional brain abnormalities between alphabetic languages varying in OD were objectively quantified by comparing foci of under- and overactivation in dyslexic readers relative to typical readers as reported in 14 studies in DO (English) and in 14 studies in SO (Dutch, German, Italian, Swedish). The in-scanner activation tasks used in these 28 studies included silent reading, reading aloud, (phonological) lexical decision, rhyme judgment, semantic judgment, and sentence comprehension. Importantly, the two sets of studies in DO and SO, respectively, were balanced regarding the number of tasks that explicitly required phonological processing. For an in-depth discussion on the effects of task nature and task difficulty – which are difficult to control for in coordinate-based meta-analyses – we refer to the original publication (Martin et al., 2016).

As predicted from the cross-language literature (Paulesu et al., 2001), universal reading-related dyslexic underactivation was identified in the left OT cortex including the fusiform gyrus (FFG), inferior occipital gyrus (IOG), ITG, and MTG. Specifically, eight of 14 and nine of 14 studies contributed one or more activation foci in this region for DO and SO, respectively. The large left posterior cluster of overlapping underactivation in both DO and SO relative to typical readers also included the posterior-to-anterior gradient of the visual word form system (Dehaene and Cohen, 2011; Taylor et al., 2019). These regions can be regarded as the most consistently reported regions of dyslexic underactivation relative to typical readers in alphabetic orthographies – irrespective of OD, in-scanner activation task, and age of participants (the mean age of the participants in the 28 included studies ranged from 8 to 30 years).

The direct statistical comparison between the two sets of fMRI studies revealed higher convergence of dyslexic underactivation relative to typical readers for DO compared with SO in the bilateral inferior parietal cortex. Interestingly, this abnormality was no longer found when foci reported with stronger dyslexic task-negative activation (i.e., task-related deactivation relative to the resting baseline) were not included in the meta-analysis. Furthermore, higher convergence of dyslexic underactivation relative to typical readers for DO compared with SO was found in the triangular part of the left inferior frontal gyrus (IFG), the left precuneus, and the right STG. Higher convergence of dyslexic overactivation relative to typical readers was identified in the left anterior insula.

Higher convergence of dyslexic underactivation for SO compared with DO was identified in the left FFG, left TP cortex, the orbital part of the left IFG, and left frontal operculum. On the contrary, higher convergence of dyslexic overactivation relative to typical readers was found in the left precentral gyrus. In sum, the findings are in line with the view of a biological unity of developmental dyslexia – with a core deficit in the left OT cortex and additional orthography-specific variations. Different patterns of reading-related dyslexic overactivation are assumed to reflect different compensatory mechanisms across languages.

The results of the meta-analysis by Martin et al. (2016) are summarized in **Table 1**.

Importantly, common dyslexic underactivation in alphabetic orthographies was found in the left OT cortex, including the visual word form system. The universal left OT cortex dysfunction, most probably reflecting the phonological speed deficit characteristic of developmental dyslexia, is in line with evidence showing that in typical readers this area subserves both lexical whole-word recognition and sublexical serial decoding (e.g., Richlan et al., 2010; Schurz et al., 2010; Wimmer et al., 2010; Schuster et al., 2016) – at least in the studied alphabetic orthographies.

THE NEUROBIOLOGY OF DEVELOPMENTAL DYSLEXIA IN ALPHABETIC, SYLLABIC, AND LOGOGRAPHIC WRITING SYSTEMS

In addition to the functional neuroimaging studies on reading and dyslexia in alphabetic orthographies, there have been studies on reading in syllabic (e.g., Japanese Kana), morpho-syllabic (e.g., Japanese Kanji), and logographic (e.g., Chinese) writing systems. In their meta-analysis of these studies on typical readers, Bolger et al. (2005) identified convergent reading-related activation in all of the above writing systems in a core network of the left STG, IFG, and OT cortex. A similar network of brain regions was found to show common activation across reading in Spanish, English, Hebrew, and Chinese (Rueckl et al., 2015). Accordingly, the brain activation abnormalities exhibited by dyslexic readers can probably be expected in similar regions across all writing systems. Direct evidence for this expectation, however, is still scarce.

The separate reading-related activation patterns of the different writing systems also varied to a certain extent, particularly regarding the spatial configuration of the activation clusters. Specifically, the meta-analysis by Bolger et al. (2005) identified divergence in the left STG (with more consistent activation for alphabetic and syllabic writing systems), and in the left IFG and right OT cortex (with more consistent activation for Chinese). The stronger activation for the alphabetic and syllabic writing systems in the left STG was ascribed to the fact that the written symbols are mapped to more fine-grained speech sounds (phonemes and syllables), as opposed to whole-word phonology in Japanese Kanji and Chinese. The stronger activation for Chinese in the left IFG was associated with higher demands on integrated processing of semantic and phonological information, which is required for unambiguous word recognition due to the high number of homophones in Chinese.

The first evidence for a specific brain dysfunction in Chinese dyslexic reading that was previously not reported for alphabetic writing systems was put forward by Siok et al. (2004). Their fMRI study found significant dyslexic underactivation in the left middle frontal gyrus (MFG) in Chinese children during both homophone judgment and lexical decision tasks. Accordingly, it was argued by the authors that fluent Chinese reading relies on

TABLE 1 | Brain abnormalities in developmental dyslexia identified in representative studies.

Region	Functional abnormalities in alphabetic orthographies (Martin et al., 2016)	Functional abnormalities in syllabic/ logographic writing systems	Structural abnormalities
L occipitotemporal cortex	Underactivation in both deep and shallow orthographies	Underactivation in Chinese (Siok et al., 2004; Hu et al., 2010)	
B inferior parietal lobule	Higher underactivation in deep (due to stronger deactivation)		
L inferior frontal gyrus, triangular	Higher underactivation in deep	Overactivation in Chinese (Siok et al., 2004)	
L precuneus	Higher underactivation in deep		
R superior temporal gyrus	Higher underactivation in deep		Reduced gray matter volume in alphabetic (Richlan et al., 2013)
L anterior insula	Higher overactivation in deep		
L fusiform gyrus	Higher underactivation in shallow	Underactivation in Chinese (Siok et al., 2004; Hu et al., 2010)	
L temporoparietal cortex	Higher underactivation in shallow	Underactivation in Chinese (Hu et al., 2010)	Reduced gray matter volume in alphabetic (Richlan et al., 2013; Eckert et al., 2016)
L inferior frontal gyrus, orbital	Higher underactivation in shallow		Reduced gray matter volume in alphabetic (Eckert et al., 2016)
L frontal operculum	Higher underactivation in shallow	Underactivation in Chinese (Siok et al., 2004)	
L precentral gyrus	Higher overactivation in shallow		
L middle frontal gyrus		Underactivation in Chinese (Siok et al., 2004; Hu et al., 2010)	Reduced gray matter volume in Chinese (Siok et al., 2008)
L dorsal inferior frontal gyrus		Underactivation in Chinese (Cao et al., 2017)	
R cerebellum			Reduced gray matter volume in alphabetic (Eckert et al., 2016)

the integrity of the left MFG as a main hub for the coordination and integration of information in verbal and spatial working memory and that developmental dyslexia results from a failure of this brain region (Perfetti et al., 2006).

The left MFG was also identified in a direct cross-linguistic comparison between dyslexic and typical readers of Chinese and English using a semantic word matching task (Hu et al., 2010). Despite brain activation differences between Chinese and English typical readers, the dyslexic readers of both writing systems showed a similar pattern of underactivation compared with the typical readers in the left MFG, left TP cortex, and left OT cortex. That is, in contrast to previous studies (see **Table 1**), even the English dyslexic readers were identified as exhibiting underactivation in the left MFG. Therefore, the functional neuroanatomical signature of developmental dyslexia in Chinese and English seems to be more similar than originally proposed by Siok et al. (2004) and reflected in underactivation of a common network including left (middle) frontal, TP, and OT regions – at least when a semantic processing task is used during brain scanning.

A remarkably similar brain network was identified by Cao et al. (2017) using an auditory rhyme judgment task. Specifically, they found that Chinese children with developmental dyslexia exhibited underactivation of a left dorsal IFG region relative to both age-matched and reading performance-matched control participants. Although anatomically labeled as left IFG, the maximum of the activation cluster was in close proximity to the left MFG with an Euclidean distance of only 16 mm and

8 mm to the peaks reported by Hu et al. (2010) and Siok et al. (2004), respectively. This left IFG dysfunction was associated with a phonological processing deficit of dyslexic readers that correlated with the severity of reading problems. Furthermore, analyses of functional connectivity identified weaker connections between the left IFG and left FFG and between the left STG and left FFG in dyslexic readers compared with the control participants. These findings were interpreted as reflecting a problem in the connection of orthography and phonology in Chinese developmental dyslexia.

STRUCTURAL BRAIN ABNORMALITIES IN DEVELOPMENTAL DYSLEXIA ACROSS LANGUAGES

Seminal neurological examinations on the neural basis of acquired reading problems were already conducted in the nineteenth century (Dejerine, 1891, 1892). In the case of developmental reading problems, neurological studies in the 1970s and 1980s were based on histological brain examinations. For example, Galaburda and Kemper (1979) identified reduced left-right asymmetry of the planum temporale in a post-mortem brain examination of a dyslexic reader. Further studies by Galaburda et al. (1985) and Humphreys et al. (1990) reported additional structural abnormalities such as neuronal ectopias and architectural dysplasias in the left TP cortex of four more dyslexia cases.

The advent of modern-day neuroimaging technology and the development of Voxel-Based Morphometry (VBM; Ashburner and Friston, 2000), enabled the automatic and objective analysis of brain structure *in vivo*. In short, VBM provides a measure of local GM volume or density of a voxel. It is an established method in cognitive neuroscience and has been used to investigate pre-reading children with a familial risk for dyslexia (e.g., Raschle et al., 2011, 2015; Black et al., 2012), dyslexic children (e.g., Eckert et al., 2005; Hoeft et al., 2007; Kronbichler et al., 2008; Krafnick et al., 2014; Jednoróg et al., 2015), and dyslexic adults (e.g., Brown et al., 2001; Brambati et al., 2004; Silani et al., 2005; Steinbrink et al., 2008; Pernet et al., 2009).

Regarding structural abnormalities in the brain of Chinese dyslexic readers, first evidence was again reported by Siok et al. (2008). Similar to the region identified with dyslexic underactivation in their previous functional MRI study (Siok et al., 2004), they found reduced GM volume in the left MFG of dyslexic children. Crucially, no other cortical or subcortical regions exhibited differences in GM volume between dyslexic and typical readers of Chinese, even in a sensitive regions-of-interest analysis focused on the left MTG, TP, and OT cortex.

A recent study (Qi et al., 2016) examined large-scale brain networks in Chinese dyslexic children. In their analysis of structural T1-weighted MRI data they distinguished between two complementary measurements of neuroanatomy in order to disentangle early congenital effects from later developed effects. Specifically, whereas the measurement of cortical surface area is thought to be sensitive to prenatal development, the measurement of cortical thickness is thought to be more sensitive to postnatal development. The Chinese dyslexic children exhibited abnormalities in both measurements, in the sense that the structural brain networks of the dyslexic children were more bilateral (i.e., less lateralized), more distributed in anterior brain regions, and less distributed in posterior brain regions compared with the typically reading children.

Due to the substantial number of existing VBM studies on dyslexia, objective coordinate-based meta-analyses were used in order to identify and specify stable effects across studies (e.g., Richlan et al., 2013). As shown in **Table 1**, consistent GM volume reduction in developmental dyslexia in alphabetic orthographies was identified in the right STG and in the left superior temporal sulcus (STS). The robustness of these findings, however, was limited as convergence across studies was relatively weak with only about half of the studies contributing to the meta-analytic clusters.

The limited convergence across studies was recently critically examined in more detail by Ramus et al. (2018). They argued that most VBM studies on developmental dyslexia are based on relatively few and relatively heterogeneous participants, leading to a high number of false positive rates in the primary literature and, therefore, little replicability of results across independent studies. This issue concerns cross-linguistic comparisons probably even more, with additional sources of heterogeneity including different assessment tools, educational systems, and socio-demographic factors.

Nevertheless, the findings of our meta-analysis found plausible support in other structural neuroanatomical studies on reading and dyslexia. The right STG region was a focal point in a remarkable and unique study by Carreiras et al. (2009). In this study, the researchers investigated (ex-) illiterates who did (or did not) learn to read as adults. The main finding was that learning to read was accompanied by an increase in GM volume in bilateral TP and dorsal occipital regions. Concerning the meta-analysis on structural brain abnormalities in developmental dyslexia, this result indicates that the right STG GM volume reduction exhibited by dyslexic readers might reflect their reduced reading experience. Therefore, the GM volume reduction is a consequence rather than a cause of reading problems in developmental dyslexia.

Two VBM studies with pre-reading children, however, support a different interpretation of the right STG GM volume reduction. Specifically, Raschle et al. (2011), reported that children with a high family-risk for developmental dyslexia were identified as having reduced GM volume in both left and right TP cortex even before formal reading instruction. Likewise, Black et al. (2012) found that a family history of reading disability was related to a reduction in GM volume in the bilateral TP cortex of five to 6-year old beginner readers. In this age group, the structural brain abnormalities can hardly be explained by a reduced amount of reading experience.

While the GM volume reduction in the right STG was an unexpected finding of our meta-analysis, the GM volume reduction in the left STS was not. The left STS GM volume reduction is in line with a large body of evidence for left perisylvian cortical anomalies in dyslexia, as identified in the already mentioned post-mortem brain examinations (e.g., Galaburda et al., 1985) and in early brain imaging studies (Eliez et al., 2000). Crucially, a similar left temporal region was identified as showing GM volume reduction across Italian, French, and English adult dyslexic readers (Silani et al., 2005). More recently, the left STS was shown to be one of the most reliable regions identified with reduced GM volume in developmental dyslexia in a combined meta-analysis and multi-center study across different laboratories from the United States (Eckert et al., 2016).

In order to interpret the functional effect of left STS abnormality in developmental dyslexia, it is important to investigate its role in typical and disrupted language processing. Classically, neurological lesions of the left STS were linked to problems in speech comprehension (Wernicke's aphasia). In more up-to-date conceptions on the neurology of speech and language (e.g., Hickok and Poeppel, 2007), the function of the left STS is associated with the representation and processing of multimodal phonological information. Therefore, it is recruited by both perceptual and productive speech processes, as well as by working memory processes involving phonological information. These cognitive functions are particularly crucial for a successful start at the beginning of literacy acquisition across languages.

Across different alphabetic orthographies, the left STS is assumed to play an important role in the integration of auditory and visual information (e.g., van Atteveldt et al., 2004; Blomert, 2011; Holloway et al., 2013; Richlan, 2019). Therefore, during skilled reading and especially during typical reading acquisition, it is recruited by self-reliant learning processes based on serial

grapheme-to-phoneme conversion. The structural GM volume reduction in the left STS in developmental dyslexia might be related to problems in this sublexical self-teaching reading strategy. Specifically, it was proposed that dyslexic readers suffer from a disruption in the development of a brain system for efficient interactive processing of auditory and visual linguistic inputs (Blau et al., 2010). Taken together, the existing evidence suggests that left STS and right STG GM volume reductions are reliable neuroanatomical signatures of adult dyslexia across different alphabetic orthographies, which might exist even before the onset of formal reading instruction.

LIMITATIONS AND FUTURE DIRECTIONS

Cross-linguistic comparisons have proven to provide extremely valuable information on the neurobiology of reading and developmental dyslexia. The focus, up to now, was largely on the comparison of dysfunctions in the form of reading-related dyslexic underactivation relative to typical readers. In contrast, the patterns of dyslexic overactivation relative to typical readers were rarely compared across languages and writing systems. This is probably because there is larger inter-individual variability with respect to overactivation compared with underactivation in developmental dyslexia and, in turn, less consistency across studies (and activation tasks). From the results reported by Martin et al. (2016), it seems that OD plays a role in the consistency of dyslexic overactivation patterns, with English dyslexic readers exhibiting more heterogeneous patterns compared with dyslexic readers from SO. This leads to only a single meta-analytic cluster identified with overactivation in English dyslexic readers compared with seven meta-analytic clusters in dyslexic readers from SO.

In principle, the dyslexic overactivation patterns might be informative on potential compensatory mechanisms supporting language-specific or language-universal remediation strategies. First evidence (Martin et al., 2016; Cao et al., 2017; Hancock et al., 2017) points to an important role of the precentral gyrus possibly subserving such neural compensation. At least in alphabetic orthographies, this compensatory role was attributed to increased reliance on articulatory processing in dyslexic readers (Hancock et al., 2017), particularly for dyslexic readers from SO. Future studies across different languages and writing systems, however, are urgently needed to shed more light on this issue.

One way of providing this kind of evidence is via intervention studies and longitudinal research. These longitudinal brain imaging studies would also be helpful for a better understanding of the relationship between brain function and brain structure and their respective effects on reading development across languages. Unfortunately, such cross-linguistic longitudinal studies are extremely challenging to conduct and to analyze,

and therefore, do not exist yet. Certainly, more fundamental research on the interplay between the developmental changes in brain function, brain structure and literacy acquisition is required in order to put forward comprehensive brain-based models of typical and dyslexic reading development.

CONCLUSION

Across alphabetic writing systems, OD has an influence on the relative importance of different underlying cognitive processes required for fluent reading, and accordingly on the degree and spatial extent of brain activation clusters of typical readers. Consequently, the neuroanatomical dysfunctions of dyslexic readers are associated with an emphasis on different elements of the core reading network, reflected in stronger or weaker under- and overactivation relative to typical readers depending on OD. For example, in the case of the logographic Chinese writing system, a crucial role is assigned to the left MFG, which possibly subserves the working memory processes required for the successful recognition of written characters.

The existing evidence, up to now, suggests that the functional neuroanatomy of developmental dyslexia is similar across languages and writing systems, with some orthography-specific peculiarities. Specifically, underactivation (in dyslexic readers relative to typical readers) in core regions of the left hemisphere reading network including OT, TP, and IFG regions in response to reading or reading-related tasks seems to be a universal signature of developmental dyslexia. At least parts of the core network were also identified with structural neuroanatomical abnormalities in dyslexic readers – sometimes even before the onset of formal reading instruction (in children with a familial risk for developmental dyslexia). Consequently, these core regions are language-universal prime candidates to be targeted by intervention programs.

REFERENCES

Alvarez, T. A., and Fiez, J. A. (2018). Current perspectives on the cerebellum and reading development. *Neurosci. Biobehav. Rev.* 92, 55–66. doi: 10.1016/j.neubiorev.2018.05.006

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Washington, DC: Author.

Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821. doi: 10.1006/nimg.2000.0582

Ben-Shachar, M., Dougherty, R. F., and Wandell, B. A. (2007). White matter pathways in reading. *Curr. Opin. Neurobiol.* 17, 258–270. doi: 10.1016/j.conb. 2007.03.006

Black, J. M., Tanaka, H., Stanley, L., Nagamine, M., Zakerani, N., Thurston, A., et al. (2012). Maternal history of reading difficulty is associated with reduced language-related gray matter in beginning readers. *Neuroimage* 59, 3021–3032. doi: 10.1016/j.neuroimage.2011.10.024

Blau, V., Reithler, J., Van Atteveldt, N., Seitz, J., Ferretsen, P., Goebel, R., et al. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain* 133, 868–879. doi: 10.1093/brain/awp308

Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *Neuroimage* 57, 695–703. doi: 10.1016/j.neuroimage.2010.11.003

Bolger, D. J., Perfetti, C. A., and Schneider, W. (2005). Cross-cultural effect on the brain revisited: universal structures plus writing system variation. *Hum. Brain Mapp.* 25, 92–104. doi: 10.1002/hbm.20124

Brambati, S. M., Termine, C., Ruffino, M., Stella, G., Fazio, F., Cappa, S. F., et al. (2004). Regional reductions of gray matter volume in familial dyslexia. *Neurology* 53, 742–745. doi: 10.1212/01.wnl.0000134673.95020.ee

Brown, W. E., Eliez, S., Menon, V., Rumsey, J. M., White, C. D., and Reiss, A. L. (2001). Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology* 56, 781–783. doi: 10.1212/wnl.56.6.781

Cao, F., Bitan, T., and Booth, J. R. (2008). Effective brain connectivity in children with reading difficulties during phonological processing. *Brain Lang.* 107, 91–101. doi: 10.1016/j.bandl.2007.12.009

Cao, F., Yan, X., Wang, Z., Liu, Y., Wang, J., Spray, G. J., et al. (2017). Neural signatures of phonological deficits in Chinese developmental dyslexia. *Neuroimage* 146, 301–311. doi: 10.1016/j.neuroimage.2016.11.051

Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., et al. (2009). An anatomical signature for literacy. *Nature* 461:983. doi: 10.1038/nature08461

Catts, H. W., Petscher, Y., Schatschneider, C., Sittner, B. M., and Mendoza, K. (2008). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *J. Learn. Disabil.* 42, 163–176. doi: 10.1177/0022219408326219

Dehaene, S., and Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends Cogn. Sci.* 15, 254–262. doi: 10.1016/j.tics.2011.04.003

Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16, 234–244. doi: 10.1038/nrn3924

Dejerine, J. (1891). Sur un cas de cécité verbale avec agraphie, suivi d'autopsie. *Mem. Soc. Biol.* 3, 197–201.

Dejerine, J. (1892). Contribution de l'étude anatomo-pathologique et clinique des differentes varietiés de cecité verbale. *Mem. Soc. Biol.* 4, 61–90.

Eckert, M. A., Berninger, V. W., Vaden, K. I., Gebregziabher, M., Tsu, L., and Dyslexia Data Consortium, (2016). Gray matter features of reading disability: a combined meta-analytic and direct analysis approach. *eNeuro* 3:ENEURO.103-15. doi: 10.1523/ENEURO.0103-15.2015

Eckert, M. A., Leonard, C. M., Wilke, M., Eckert, M., Richards, T., Richards, A., et al. (2005). Anatomical signatures of dyslexia in children: unique information from manual and voxel based morphometry brain measures. *Cortex* 41, 304–315. doi: 10.1016/s0010-9452(08)70268-5

Eliez, S., Rumsey, J. M., Giedd, J. N., Schmitt, J. E., Patwardhan, A. J., and Reiss, A. L. (2000). Morphological alteration of temporal lobe gray matter in dyslexia: an MRI study. *J. Child Psychol. Psychiatry Allied Discipl.* 41, 637–644. doi: 10.1111/1469-7610.00650

Frost, R., Katz, L., and Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: a multilingual comparison. *J. Exp. Psychol.* 13:104. doi: 10.1037/0096-1523.13.1.104

Galaburda, A. M., and Kemper, T. L. (1979). Cytoarchitectonic abnormalities in developmental dyslexia: a case study. *Ann. Neurol.* 6, 94–100. doi: 10.1002/ana. 410060203

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F., and Geschwind, N. (1985). Developmental dyslexia: four consecutive patients with cortical anomalies. *Ann. Neurol.* 18, 222–233. doi: 10.1002/ana.410180210

Hancock, R., Richlan, F., and Hoeft, F. (2017). Possible roles for fronto-striatal circuits in reading disorder. *Neurosci. Biobehav. Rev.* 72, 243–260. doi: 10.1016/j.neubiorev.2016.10.025

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8:393. doi: 10.1038/nrn2113

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. K., et al. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4234–4239. doi: 10.1073/pnas. 0609399104

Holloway, I. D., van Atteveldt, N., Blomert, L., and Ansari, D. (2013). Orthographic dependency in the neural correlates of reading: evidence from audiovisual integration in English readers. *Cereb. Cortex* 25, 1544–1553. doi: 10.1093/cercor/bht347

Hu, W., Lee, H. L., Zhang, Q., Liu, T., Geng, L. B., Seghier, M. L., et al. (2010). Developmental dyslexia in Chinese and English populations: dissociating the effect of dyslexia from language differences. *Brain* 133, 1694–1706. doi: 10.1093/brain/awq106

Humphreys, P., Kaufmann, W. E., and Galaburda, A. M. (1990). Developmental dyslexia in women: neuropathological findings in three patients. *Ann. Neurol.* 28, 727–738. doi: 10.1002/ana.410280602

Jednoróg, K., Marchewka, A., Altarelli, I., Monzalvo Lopez, A. K., van Ermingen-Marbach, M., Grande, M., et al. (2015). How reliable are gray matter disruptions in specific reading disability across multiple countries and languages? Insights from a large-scale voxel-based morphometry study. *Hum. Brain Mapp.* 36, 1741–1754. doi: 10.1002/hbm.22734

Koyama, M. S., Di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz, S. J., et al. (2013). Cortical signatures of dyslexia and remediation: an intrinsic functional connectivity approach. *PLoS One* 8:e55454. doi: 10.1371/journal.pone.0055454

Krafnick, A. J., Flowers, D. L., Luetje, M. M., Napoliello, E. M., and Eden, G. F. (2014). An investigation into the origin of anatomical differences in dyslexia. *J. Neurosci.* 34, 901–908. doi: 10.1523/JNEUROSCI.2092-13.2013

Kronbichler, M., Wimmer, H., Staffen, W., Hutzler, F., Mair, A., and Ladurner, G. (2008). Developmental dyslexia: gray matter abnormalities in the occipitotemporal cortex. *Hum. Brain Mapp.* 29, 613–625. doi: 10.1002/hbm. 20425

Landerl, K., and Moll, K. (2010). Comorbidity of learning disorders: prevalence and familial transmission. *J. Child Psychol. Psychiatry* 51, 287–294. doi: 10.1111/j. 1469-7610.2009.02164.x

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H., Lohvansuu, K., et al. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *J. Child Psychol. Psychiatry* 54, 686–694. doi: 10.1111/jcpp.12029

Lyon, R. G., Shaywitz, S. E., and Shaywitz, B. A. (2003). A definition of dyslexia. *Ann. Dyslexia* 53, 1–14. doi: 10.1007/s11881-003-0001-9

Martin, A., Kronbichler, M., and Richlan, F. (2016). Dyslexic brain activation abnormalities in deep and shallow orthographies: a meta-analysis of 28 functional neuroimaging studies. *Hum. Brain Mapp.* 37, 2676–2699. doi: 10. 1002/hbm.23202

Martin, A., Schurz, M., Kronbichler, M., and Richlan, F. (2015). Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 36, 1963–1981. doi: 10.1002/hbm.22749

Olulade, O. A., Flowers, D. L., Napoliello, E. M., and Eden, G. F. (2015). Dyslexic children lack word selectivity gradients in occipito-temporal and inferior frontal cortex. *Neuroimage* 7, 742–754. doi: 10.1016/j.nicl.2015.02.013

Paulesu, E., Demonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: cultural diversity and biological unity. *Science* 291, 2165–2167. doi: 10.1126/science.1057179

Perfetti, C. A., Tan, L. H., and Siok, W. T. (2006). Brain-behavior relations in reading and dyslexia: implications of Chinese results. *Brain Lang.* 98, 344–346. doi: 10.1016/j.bandl.2006.04.010

Pernet, C., Andersson, J., Paulesu, E., and Demonet, J. F. (2009). When all hypotheses are right: a multifocal account of dyslexia. *Hum. Brain Mapp.* 30, 2278–2292. doi: 10.1002/hbm.20670

Peterson, R. L., Pennington, B. F., Shriberg, L. D., and Boada, R. (2009). What influences literacy outcome in children with speech sound disorder? *J. Speech Lang. Hear. Res.* 52, 1175–1188. doi: 10.1044/1092-4388(2009/08-0024)

Pugh, K. R., Sandak, R., Frost, S. J., Moore, D., and Mencl, W. E. (2005). Examining reading development and reading disability in English language learners: potential contributions from functional neuroimaging. *Learn. Disabil. Res. Pract.* 20, 24–30. doi: 10.1111/j.1540-5826.2005.00117.x

Qi, T., Gu, B., Ding, G., Gong, G., Lu, C., Peng, D., et al. (2016). More bilateral, more anterior: alterations of brain organization in the large-scale structural network in Chinese dyslexia. *Neuroimage* 124, 63–74. doi: 10.1016/j.neuroimage.2015.09.011

Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., and di Covella, L. S. (2018). Neuroanatomy of developmental dyslexia: pitfalls and promise. *Neurosci. Biobehav. Rev.* 84, 434–452. doi: 10.1016/j.neubiorev.2017.08.001

Raschle, N. M., Becker, B. L. C., Smith, S., Fehlbaum, L. V., Wang, Y., and Gaab, N. (2015). Investigating the influences of language delay and/or familial risk for dyslexia on brain structure in 5-year-olds. *Cereb. Cortex* 27, 764–776. doi: 10.1093/cercor/bhv267

Raschle, N. M., Chang, M., and Gaab, N. (2011). Structural brain alterations associated with dyslexia predate reading onset. *Neuroimage* 57, 742–749. doi: 10.1016/j.neuroimage.2010.09.055

Richlan, F. (2012). Developmental dyslexia: dysfunction of a left hemisphere reading network. *Front. Hum. Neurosci.* 6:120. doi: 10.3389/fnhum.2012.00120

Richlan, F. (2014). Functional neuroanatomy of developmental dyslexia: the role of orthographic depth. *Front. Hum. Neurosci.* 8:347. doi: 10.3389/fnhum.2014.00347

Richlan, F. (2019). The functional neuroanatomy of letter-speech sound integration and its relation to brain abnormalities in developmental dyslexia. *Front. Hum. Neurosci.* 13:21. doi: 10.3389/fnhum.2019.00021

Richlan, F., Kronbichler, M., and Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: a quantitative meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* 30, 3299–3308. doi: 10.1002/hbm.20752

Richlan, F., Kronbichler, M., and Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage* 56, 1735–1742. doi: 10.1016/j.neuroimage.2011.02.040

Richlan, F., Kronbichler, M., and Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: a meta-analysis of Voxel-Based Morphometry studies. *Hum. Brain Mapp.* 34, 3055–3065. doi: 10.1002/hbm.22127

Richlan, F., Sturm, D., Schurz, M., Kronbichler, M., Ladurner, G., and Wimmer, H. (2010). A common left occipito-temporal dysfunction in developmental dyslexia and acquired letter-by-letter reading? *PLoS One* 5:e12073. doi: 10.1371/journal.pone.0012073

Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W. J., Bick, A., Frost, S. J., et al. (2015). Universal brain signature of proficient reading: evidence from four contrasting languages. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15510–15515. doi: 10.1073/pnas.1509321112

Schurz, M., Sturm, D., Richlan, F., Kronbichler, M., Ladurner, G., and Wimmer, H. (2010). A dual-route perspective on brain activation in response to visual words: evidence for a length by lexicality interaction in the visual word form area (VWFA). *Neuroimage* 49, 2649–2661. doi: 10.1016/j.neuroimage.2009.10.082

Schurz, M., Wimmer, H., Richlan, F., Ludersdorfer, P., Klackl, J., and Kronbichler, M. (2015). Resting-state and task-based functional brain connectivity in developmental dyslexia. *Cereb. Cortex* 25, 3502–3514. doi: 10.1093/cercor/bhu184

Schuster, S., Hawelka, S., Hutzler, F., Kronbichler, M., and Richlan, F. (2016). Words in context: the effects of length, frequency and predictability on brain responses during natural reading. *Cereb. Cortex* 26, 3889–3904. doi: 10.1093/cercor/bhw184

Seymour, P. H., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859

Silani, G., Frith, U., Demonet, J. F., Fazio, F., Perani, D., Price, C., et al. (2005). Brain abnormalities underlying altered activation in dyslexia: a voxel based morphometry study. *Brain* 128, 2453–2461. doi: 10.1093/brain/awh579

Siok, W. T., Niu, Z., Jin, Z., Perfetti, C. A., and Tan, L. H. (2008). A structural–functional basis for dyslexia in the cortex of Chinese readers. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5561–5566. doi: 10.1073/pnas.0801750105

Siok, W. T., Perfetti, C. A., Jin, Z., and Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature* 431, 71–76. doi: 10.1038/nature02865

Steinbrink, C., Vogt, K., Kastrup, A., Müller, H. P., Juengling, F. D., Kassubek, J., et al. (2008). The contribution of white and gray matter differences to developmental dyslexia: insights from DTI and VBM at 3.0. *Neuropsychologia* 46, 3170–3178. doi: 10.1016/j.neuropsychologia.2008.07.015

Taylor, J. S. H., Davis, M. H., and Rastle, K. (2019). Mapping visual symbols onto spoken language along the ventral visual stream. *Proc. Natl. Acad. Sci. U.S.A.* 116, 201818575. doi: 10.1073/pnas.1818575116

van Atteveldt, N., Formisano, E., Goebel, R., and Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron* 43, 271–282. doi: 10.1016/j.neuron.2004.06.025

van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., et al. (2011). The left occipitotemporal system in reading: disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. *Neuroimage* 54, 2426–2436. doi: 10.1016/j.neuroimage.2010.10.002

Vandermosten, M., Boets, B., Wouters, J., and Ghesquière, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neurosci. Biobehav. Rev.* 36, 1532–1552. doi: 10.1016/j.neubiorev.2012.04.002

Willcutt, E. G., Betjemann, R. S., McGrath, L. M., Chhabildas, N. A., Olson, R. K., DeFries, J. C., et al. (2010). Etiology and neuropsychology of comorbidity between RD and ADHD: the case for multiple-deficit models. *Cortex* 46, 1345–1361. doi: 10.1016/j.cortex.2010.06.009

Wimmer, H., Schurz, M., Sturm, D., Richlan, F., Klackl, J., Kronbichler, M., et al. (2010). A dual-route perspective on poor reading in a regular orthography: an fMRI study. *Cortex* 46, 1284–1298. doi: 10.1016/j.cortex.2010.06.004

World Health Organization (2016). *ICD-10 – International Statistical Classification of Diseases and Related Health Problems 10th Revision*, 5th Edn. Geneva: World Health Organization.

The Overlap of Poor Reading Comprehension in English and French

Nadia D'Angelo[1]*, Klaudia Krenca[2] and Xi Chen[2]

[1] Ontario Ministry of Education, Toronto, ON, Canada, [2] Department of Applied Psychology & Human Development, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

This study examined overlap and correlates of poor reading comprehension in English and French for children in early French immersion. Poor comprehenders were identified in grade 3 in English and French using a regression method to predict reading comprehension scores from age, non-verbal reasoning, word reading accuracy, and word reading fluency. Three groups of poor comprehenders were identified: 10 poor comprehenders in English and French, 11 poor comprehenders in English, and 10 poor comprehenders in French, and compared to 10 controls with good reading comprehension in both English and French. There was a moderate degree of overlap in comprehension difficulties in English and French among poor comprehenders with equivalent amounts of exposure to French, with a prevalence rate of 41.7% in our sample. Children who were poor comprehenders in both English and French consistently scored the lowest on English vocabulary in grade 1 and grade 3 and in French vocabulary in grade 3 suggesting that poor comprehenders' vocabulary weaknesses in English as a primary language may contribute to comprehension difficulties in English and French.

Keywords: poor comprehenders, reading comprehension, French immersion, oral language skills, vocabulary, comprehension difficulties, bilingual learners

INTRODUCTION

There is considerable evidence to suggest that children who are at risk for reading difficulties in a second language (L2) can be identified through early assessment of word reading and cognitive skills in their first language (L1), before their oral language proficiency is fully developed in L2 (Geva and Clifton, 1994; Da Fontoura and Siegel, 1995; MacCoubrey et al., 2004). Much of this previous research is based on the premise that certain cognitive and linguistic skills, such as phonological processing, transfer across languages (e.g., Comeau et al., 1999; August and Shanahan, 2006). More recently, studies have investigated children's reading comprehension difficulties that occur despite age-appropriate decoding skills (e.g., Nation et al., 2010; Tong et al., 2011). Relatively little is known about the identification of poor reading comprehension in the absence of poor decoding, and even less is known about whether reading comprehension difficulties manifest in a similar manner in L1 and L2 for children learning in a bilingual context. The present study aims to investigate overlap and early contributors of poor reading comprehension for children in early French immersion programs in Canada who receive school instruction in French, an additional language, while being exposed to English, their primary language of the community.

Reading comprehension is a complex process that involves the integration and coordination of various skills, including word decoding, the ability to decipher or recognize printed words, and oral language or listening comprehension, the ability to understand what is decoded in spoken form (Simple View of Reading; Gough and Tunmer, 1986). Most research into reading comprehension difficulties has focused on children with poor decoding whose weaknesses manifest early in reading development as phonological awareness and word reading deficits (e.g., Snowling, 2000). In contrast to poor decoders, poor comprehenders' difficulties appear to emerge later, when decoding becomes automatized and more variance in reading comprehension is accounted for by oral language skills (Catts et al., 2012). Oral language difficulties tend to be masked by poor comprehenders' age-appropriate decoding skills, and as a result, early indicators of later reading comprehension difficulties are often overlooked.

Existing longitudinal studies have used a retrospective approach to examine poor comprehenders' deficits across previous grades and suggest that oral language weaknesses are prevalent in poor comprehenders before their reading comprehension difficulties become apparent (Catts et al., 2006; Nation et al., 2010; Tong et al., 2011). For example, Nation et al. (2010) identified poor comprehenders based on reading achievement at age 8 and retrospectively examined their reading and language skills beginning at age 5. While poor comprehenders' phonological processing and word reading skills progressed over time, their oral language skills remained persistently weak, suggesting that early weaknesses in understanding and producing spoken language contributed to poor comprehenders' comprehension difficulties.

The linguistic interdependence hypothesis suggests that L1 and L2 reading skills are interdependent, and that language and literacy skills acquired in one language facilitate reading development in the L2 (Cummins, 1984). Thus, it seems probable that the same cognitive and linguistic skills needed for successful reading comprehension in L1 contribute to reading development in L2 (e.g., Gottardo and Mueller, 2009; Mancilla-Martinez and Lesaux, 2010). Indeed, previous research suggests that it is possible to identify children at-risk for L2 reading difficulties based on their performance in L1 (Geva and Clifton, 1994; Da Fontoura and Siegel, 1995). However, few studies have investigated poor comprehenders in a bilingual context largely due to the complexity of understanding reading comprehension processes in L1 and L2. Children learning in an L2 are in the process of acquiring the language of instruction and it may be difficult to determine whether weaknesses in L2 reading comprehension reflect limited language learning experiences or are indicative of a language or reading impairment (Paradis et al., 2010; Li and Kirby, 2014; D'Angelo and Chen, 2017).

Li and Kirby (2014) examined the reading comprehension profiles of grade 8 emerging Chinese-English bilinguals in an English immersion program in China. Poor comprehenders were distinguished from average comprehenders based on their performance on English L2 vocabulary measures. The authors concluded that because the groups did not differ on Chinese L1 word reading and reading comprehension, poor comprehenders'

reading comprehension difficulties were due to limited English L2 proficiency. However, the comprehender groups in this study were selected using English L2 assessments only and therefore, children with an underlying oral language impairment across the two languages could not be identified. Since Chinese and English and are not closely related languages, vocabulary and reading comprehension may not have the same underlying mechanisms in each language.

A few studies have identified poor comprehenders based on English L1 reading performance in a French immersion context and suggest that poor comprehenders demonstrate relatively poor oral language skills in both English L1 and French L2 (e.g., D'Angelo et al., 2014; D'Angelo and Chen, 2017). D'Angelo et al. (2014) retrospectively investigated the reading and language abilities of a small sample of English L1 children in French immersion who were identified as poor and average comprehenders based on their English L1 reading performance in grade 3. They found that poor comprehenders scored relatively lower on English and French vocabulary across grades 1 to 3, despite average phonological awareness and word reading skills in both languages. Such findings suggest that poor comprehenders may indeed have an underlying problem in oral language. The current study extends the existing research to a larger, more representative sample of children in French immersion to facilitate comparison. The purpose is to determine the extent to which those identified as having poor reading comprehension in English, the societal language, also demonstrate poor reading comprehension in French, an additional language and the language of instruction.

Studies that have examined the co-occurrence of reading difficulties between an L1 and L2 have primarily focused on poor readers and suggest that there is some overlap of reading difficulty in L1 and L2 (Manis and Lindsey, 2010; McBride-Chang et al., 2013; Tong et al., 2015; Shum et al., 2016). For example, Manis and Lindsey (2010) found that 55% of grade 5 children who met the criteria for reading difficulties in English L2 (decoding scores at or below the 25th percentile) were also identified with reading difficulties in Spanish L1. Similarly, McBride-Chang et al. (2013) tested the overlap of poor readers in Chinese L1 and English L2 (defined as those at or below the 25th percentile on Chinese and English word reading tests) among 8-year-old children in Beijing and found that 40% of poor readers in Chinese L1 were also poor readers in English L2. In each study, children who were identified as poor readers in both languages scored lower on cognitive and linguistic tasks than children who were poor readers in only one language. On the other hand, children with poor reading in one language did not necessarily have difficulties in the other. It appears that the degree of overlap between poor reading is increased when the two languages are more closely related. However, these studies focused on the overlap status of poor readers based on poor decoding. We were interested in whether such overlap occurs for poor comprehenders who show discrepancies between their reading comprehension and decoding skills.

Only one known study at this time has explored the overlap between L1 and L2 reading comprehension difficulties. Tong et al. (2017) examined the co-occurrence of reading comprehension

difficulties and associated longitudinal correlates in 10-year-old children with poor reading comprehension (defined as those at or below the 25th percentile on reading comprehension tasks) in Chinese L1 and English L2. The authors found that approximately half (53%) of children with poor reading comprehension in Chinese L1 also experienced poor reading comprehension in English L2. Results indicated that word reading and language skills were longitudinal correlates of poor reading comprehension in Chinese and English. This study was among the first to investigate overlap of reading comprehension difficulties in L1 and L2 and to retrospectively examine sources of poor reading comprehension. However, the selection method used in this study identified poor comprehenders based on reading comprehension scores only and did not distinguish between children with poor oral language skills from those with poor decoding skills. In the present study, we aimed to understand the overlap of poor reading comprehension in English and French in the absence of decoding problems.

Given the challenges associated with defining poor reading comprehension in an additional language, the goal of the present study was to extend previous research on reading comprehension difficulties to English–French bilinguals to answer two specific research questions.

First, we asked whether children identified as poor comprehenders in English are also identified as poor comprehenders in French. Whereas most previous studies have examined overlap with word reading and reading comprehension scores at or below an arbitrary cut-off score, we utilized a regression technique to identify poor comprehenders in English and French by examining associations between reading comprehension scores, age, non-verbal reasoning, word reading accuracy, and word reading fluency. This approach defines groups more precisely than the cut-off score method because it examines relative discrepancies between various skills related to reading comprehension by distinguishing poor comprehenders from average and good comprehenders (e.g., Tong et al., 2011, 2014; Li and Kirby, 2014; D'Angelo and Chen, 2017).

Second, we asked what reading and language skills distinguish between poor comprehenders in English and French, poor comprehenders in English, and poor comprehenders in French. We anticipated that children identified as poor comprehenders in both English and French would show early and persistent oral language difficulties in both languages. English and French share many similarities in vocabulary, morphology, and syntax (e.g., LeBlanc and Seguin, 1996; Roy and Labelle, 2007; D'Angelo and Chen, 2017; D'Angelo et al., 2017). Both are represented by the Roman alphabet and an opaque writing system (Seymour et al., 2003). These shared structural properties are thought to facilitate cross-language associations between two languages (Koda, 2008). Therefore, we expected to see similar characteristics of reading comprehension difficulties between the two languages.

The socio-linguistic and educational context of the current study makes it possible to assess and compare English and French reading outcomes among children acquiring both languages. In Canada, French immersion is an additive dual language program that promotes oral and written language proficiency in both English and French, the official

languages. Children in early French immersion programs are non-francophones who receive integrated language and content instruction primarily in French beginning in kindergarten or grade 1. However, these children often live in predominantly English-speaking environments with limited opportunity to hear and speak French outside of the classroom. Thus, French immersion classrooms are comprised of English-speaking children for whom French is the L2 and minority language children for whom English is the L2 and French the L3. English language arts instruction is generally introduced in grade 4.

Since the children in this study had similar and limited levels of French proficiency upon school entry, any differences in French reading and language abilities between children would be unlikely a result of differences in the amount of exposure the children had to French. Specifically, for children with poor reading comprehension in both English and French, we could be confident that weaknesses in oral language reflect a pervasive language impairment rather than a less developed French proficiency.

MATERIALS AND METHODS

Participants

Participants were 180 children consisting of 83 males and 97 females who were recruited from early French immersion schools in a large Canadian city and tested in English and French in the spring of grade 1 (M_{age} = 80.36 months, SD = 4.18) and grade 3 (M_{age} = 104.66 months, SD = 4.06). As part of the inclusion criteria, children selected for this study were non-native speakers of French receiving school instruction entirely in French since school entry. Out of the 180 children, 135 (75%) spoke English as a primary language. Forty-five children (25%) were exposed to additional languages at home.

Measures

The data in this study are from longitudinal research, in which several reading-related tasks were administered to participants between grades 1 and 3. Trained research assistants, who were fluent in the respective test language, administered tasks to participants at school. English and French instructions were used for French measures to ensure comprehension of the task. The order of the sessions was counterbalanced across participants and within each session the order of the task administration was randomized. Due to limited testing time, not all the same tasks were administered in each year of the study.

Non-verbal Reasoning

Children were administered the reasoning by analogy subtest of the Matrix Analogies Test in English to assess non-verbal reasoning in grade 1 (expanded form; Naglieri, 1985). For each item, children were asked to complete a figural matrix by choosing the missing piece from 5 to 6 possible choices. There were 16 items and testing was discontinued after four consecutive errors.

Phonological Awareness

This task was measured in grade 1 using the elision subtest of the Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999, 2013). The examiner read individual words aloud and children were asked to delete a syllable or phoneme from each word (e.g., "say *time* without saying/*m*/"). There were 34 test items presented in order of increasing difficulty. Testing was discontinued after three consecutive errors.

A parallel measure was created to assess phonological awareness in French. Twenty-six items were selected to match characteristics of the English task (i.e., syllable and phoneme deletion) and presented in order of increasing difficulty. The administration of the test was discontinued if the children made six consecutive errors.

Vocabulary

The Peabody Picture Vocabulary was used to measure English receptive vocabulary (PPVT-IV Form A; Dunn and Dunn, 2007) in grades 1 and 3. Each time a tester orally presented a target word, the child was required to point to one of four pictures that best corresponded to that word. Testing was discontinued when the child made eight or more errors in a set of 12.

The Échelle de Vocabulaire en Images Peabody (EVIP Form A; Dunn et al., 1993) was used to assess French receptive vocabulary in both grades. The examiner read a target word and the child was asked to identify the picture that best represented the word from a set of four pictures. Testing was discontinued after six errors were made on the previous eight consecutive items.

Word Reading Accuracy

Word reading accuracy in English was assessed in grades 1 and 3 with the Letter-Word Identification subtest from the Test of Achievement, Woodcock Johnson-III (WJ-III; Woodcock et al., 2001). Children were asked to read a series of 76 letters and words that were presented in order of increasing difficulty. Testing was discontinued after participants misread the six consecutive highest-numbered items on a given page.

French word reading accuracy was assessed using an experimental task (Au-Yeung et al., 2015). The test consists of 120 items arranged in 15 sets of eight words each. The children were asked to read the words accurately and fluently. Testing was discontinued when the children misread five or more words within a set of eight words. The total score represents the number of words read correctly.

Word Reading Fluency

Children's word reading fluency in English was measured by the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (TOWRE Form A; Torgesen et al., 1999) in grade 3. Children were provided with 45 s to quickly and accurately identify as many words as they could from a vertical list of 104 items. A parallel experimental measure was created to assess word reading fluency in French.

Reading Comprehension

The comprehension subtest (Level 3 Form S) of the Gates-MacGinitie Reading Tests (GMRT; MacGinitie et al., 2000)

was used to assess children's English reading comprehension in grade 3. Children were asked to read short passages and answer 48 corresponding multiple-choice questions. The score was the total number of correct answers. Level C Form 4 of the Gates-MacGinitie Reading Tests – Second Canadian Edition (MacGinitie and MacGinitie, 1992) was translated into French and administered in the same way as the English task.

RESULTS

To prepare the data for analyses, we first examined whether there was statistical support for merging the samples of children who spoke English as a primary language at home and those who were exposed to additional home languages into one sample. A Box's M test using the grades 1 and 3 measures, indicated no significant difference in variance-covariance patterns between the two language groups on English, Box's $M = 40.88$, $p = 0.09$, and French, Box's $M = 7.74$, $p = 0.99$, reading and language measures. Based on these results, the two groups were combined to create one sample. **Table 1** presents the mean raw scores, standard scores for standardized measures, standard deviations and reliability estimates for the entire sample on all English and French measures in grade 1 and grade 3.

We selected groups of comprehenders in grade 3 using separate regression techniques for English and French measures to predict children's reading comprehension scores from age, non-verbal reasoning, word reading accuracy, and word reading fluency. These variables are correlated with reading comprehension (e.g., Deacon and Kirby, 2004; Lesaux et al., 2006) and have been widely used for identifying comprehender subgroups (Li and Kirby, 2014; Tong et al., 2014; D'Angelo and Chen, 2017). Together, the predictors explained a total of 43% of the variance in English reading comprehension and 37% of the variance in French reading comprehension. The observed reading comprehension scores were plotted against the standardized predicted scores. Children below the lower 65% confidence interval of the regression line were identified as poor comprehenders and those above the upper 65% confidence interval were identified as good comprehenders. Those children who scored within the 15% confidence interval were identified as average comprehenders. Children with very poor or good word reading skills (predicted value 1 SD above or below the mean) were not selected and excluded from analyses.

Through this regression method, we identified three groups of comprehenders in English (24 poor, 24 average, and 24 good) and three groups of comprehenders in French (24 poor, 24 average, and 24 good). Sixteen children out of the 24 poor comprehenders of English and 18 children out of the 24 poor comprehenders of French identified as English-speaking.[1] The remaining children came from diverse linguistic backgrounds and were exposed to additional languages at home, including Russian, Hebrew, and Mandarin. A chi-square test of independence indicated a non-significant relationship between the children who spoke English

---

[1]For children to be classified as English-speaking, parents had to indicate that English was spoken in the home environment 50% of the time or more.

TABLE 1 | Means, standard deviations, and reliabilities for the total sample (N = 180) on English and French measures in grade 1 and grade 3.

Measure	M	SD	Cronbach's alpha
Grade 1			
Age (in months)	80.36	4.18	
Non-verbal reasoning	4.49	3.59	0.86
English phonological awareness	15.45	6.59	0.94
English phonological awareness SS	11.34	3.06	
French phonological awareness	10.92	5.57	0.92
English vocabulary	122.86	20.23	0.95
English vocabulary SS	109.58	14.38	
French vocabulary	35.25	15.86	0.96
French vocabulary SS	69.02	14.17	
English word reading accuracy	32.43	11.10	0.95
English word reading accuracy SS	111.06	19.50	
French word reading accuracy	30.23	19.15	0.97
Grade 3			
Age (in months)	104.66	4.06	
English vocabulary	147.47	16.48	0.94
English vocabulary SS	108.47	13.30	
French vocabulary	66.57	26.20	0.97
French vocabulary SS	76.00	21.88	
English word reading accuracy	51.36	9.59	0.94
English word reading accuracy SS	109.47	14.61	
French word reading accuracy	65.40	25.49	0.98
English word reading fluency	60.74	15.31	0.97
English word reading fluency SS	95.83	17.20	
French word reading fluency	55.18	14.58	0.98
English reading comprehension	26.24	10.49	0.91
English reading comprehension SS	95.04	13.67	
French reading comprehension	17.79	7.39	0.84

SS, standard score.

as a primary language at home and those who were exposed to additional languages at home within the comprehender groups identified in English, χ^2 (1, N = 72) = 3.11, p = 0.21, and in French, χ^2 (1, N = 72) = 1.01, p = 0.61. Based on these results, and given that the children exposed to additional languages met the inclusion criteria (non-native speakers of French), they were retained in the sample.

We conducted multivariate analyses of variance (MANOVAs) to confirm the reading comprehension profiles of the English comprehender groups and to determine whether poor comprehenders differed from average and good comprehenders on English and French reading-related measures in grade 1 and grade 3. As illustrated in **Table 2**, there were no significant differences between the three groups on age, non-verbal reasoning, English and French word reading accuracy, and English and French elision in grade 1 and English and French word reading accuracy and fluency in grade 3 (all ps > 0.08). However, as expected, poor comprehenders differed significantly from average (p < 0.001) and good comprehenders (p < 0.001) on English and French reading comprehension in grade 3. Poor comprehenders also differed from average (p < 0.001) and good comprehenders (p < 0.001) on English vocabulary in

grade 1 and grade 3. Similarly, French vocabulary distinguished poor comprehenders from average comprehenders in grade 1 (p < 0.05) and grade 3 (p < 0.01).

For the comprehender groups identified using French measures, there were no significant differences between poor, average, and good comprehenders on age, non-verbal reasoning, and English and French phonological awareness in grade 1. Poor comprehenders differed significantly from average and good comprehenders on grade 1 measures of English (p < 0.01) and French vocabulary (p < 0.01) and English (p < 0.001) and French word reading accuracy (p < 0.001). In grade 3, English (p < 0.05) and French vocabulary (p < 0.001), English word reading accuracy (p < 0.001), English (p < 0.001) and French word reading fluency (p < 0.001), and English (p < 0.001) and French reading comprehension (p < 0.001) distinguished poor comprehenders from average and good comprehenders (**Table 3**).

Table 4 presents the prevalence rates of the overlap between comprehender groups in English and French. Of particular interest to this study was the number of children who were identified through the regression technique as poor comprehenders for both English and French relative to the entire sample. Three subgroups of reading comprehension difficulties in the two languages were considered: 10 children who were poor comprehenders in both English and French (PCB), 11 children who were poor comprehenders in English only (PCE), and 10 children who were poor comprehenders in French only (PCF). We selected an additional 10 children from among the good comprehenders in both English and French, matched on age and gender, to serve as the control group. In this way, we could compare the three groups of comprehenders to children who had average English and French word reading skills, but good comprehension in both English and French. There were no significant differences between the four groups on age (PCB: M = 104.26, SD = 3.97; PCE: M = 105.01, SD = 4.98; PCF: M = 104.01, SD = 4.40; Control: M = 105.02, SD = 3.46) and non-verbal reasoning (PCB: M = 3.80, SD = 3.01; PCE: M = 2.82, SD = 2.40; PCF: M = 3.80, SD = 2.25; Control: M = 5.00, SD = 4.14). Chi-square results demonstrated that the chance of poor comprehenders in English also being poor comprehenders in French was significantly above the baseline level, χ^2 (1, N = 180) = 14.02, p < 0.001.

It should be noted that children identified as poor comprehenders in English only had not been selected for a comprehender status in French. Similarly, those identified as poor comprehenders in French only did not fit a comprehender group in English. Of the remaining children who were poor comprehenders identified in English, two were average comprehenders in French and one was a good comprehender in French. Of the remaining poor comprehenders identified in French, two were average comprehenders in English and two were good comprehenders English.

The next step in our analyses was to retrospectively examine the correlates of English and French reading comprehension difficulties for each of the three subgroups of poor comprehenders and the control group. We conducted separate MANOVAs, controlling for gender, for the English and

TABLE 2 | Means (standard deviations) of poor, average, and good comprehenders selected with English measures on English and French reading and language variables in grade 1 and grade 3.

Measure	Poor (n = 24)	Average (n = 24)	Good (n = 24)		
	M (SD)	M (SD)	M (SD)	F	Pairwise comparisons[a]
Grade 1					
Age (in months)	80.06 (4.14)	80.78 (4.56)	81.42 (4.65)	0.49	
Non-verbal reasoning	3.37 (2.85)	5.08 (3.57)	5.29 (3.98)	1.81	
English phonological awareness	11.32 (3.22)	12.00 (11.92)	11.92 (3.13)	0.29	
French phonological awareness	10.96 (5.72)	11.08 (4.60)	12.33 (5.99)	0.46	
English vocabulary	107.25 (20.14)	128.50 (16.08)	137.58 (16.36)	18.73***	PC < AC < GC
French vocabulary	27.50 (11.22)	40.71 (17.04)	36.79 (17.22)	4.03*	PC < AC = GC
English word reading accuracy	30.95 (10.12)	33.54 (10.58)	33.46 (10.42)	0.40	
French word reading accuracy	33.22 (18.21)	31.37 (18.11)	32.50 (18.59)	0.61	
Grade 3					
Age (in months)	104.74 (4.41)	104.05 (4.32)	105.23 (3.63)	0.50	
English vocabulary	127.25 (20.28)	152.00 (10.67)	159.37 (9.40)	33.23***	PC < AC < GC
French vocabulary	50.05 (17.88)	73.79 (25.70)	67.96 (28.32)	5.34**	PC < AC = GC
English word reading accuracy	50.53 (10.75)	53.54 (7.45)	54.92 (5.86)	1.62	
French word reading accuracy	73.30 (28.40)	74.25 (25.98)	60.67 (19.26)	2.23	
English word reading fluency	59.21 (10.91)	61.96 (10.46)	66.79 (7.90)	2.10	
French word reading fluency	53.35 (15.67)	55.96 (11.22)	56.88 (8.24)	0.54	
English reading comprehension	15.84 (3.43)	27.25 (5.06)	40.67 (3.41)	198.51***	PC < AC < GC
French reading comprehension	15.09 (4.63)	15.08 (6.50)	23.42 (8.62)	11.89***	PC = AC < GC

[a]Equal sign indicates non-significant difference, and less-than symbol indicates $p < 0.05$ or less. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE 3 | Means (standard deviations) of poor, average, and good comprehenders selected with French measures on English and French reading and language variables in grade 1 and grade 3.

Measure	Poor (n = 24)	Average (n = 24)	Good (n = 24)		
	M (SD)	M (SD)	M (SD)	F	Pairwise comparisons[a]
Grade 1					
Age (in months)	81.11 (4.83)	80.42 (3.79)	80.05 (3.99)	0.36	
Non-verbal reasoning	4.29 (3.21)	3.95 (3.12)	4.83 (3.99)	0.36	
English phonological awareness	13.33 (6.63)	15.41 (7.29)	17.78 (12.30)	2.11	
French phonological awareness	9.38 (4.47)	11.86 (6.44)	12.30 (5.41)	1.77	
English vocabulary	113.00 (20.24)	128.59 (14.92)	128.09 (14.96)	4.75**	PC < AC < GC
French vocabulary	28.14 (10.20)	39.77 (19.25)	44.22 (18.76)	5.35**	PC < AC = GC
English word reading accuracy	25.38 (9.12)	35.64 (8.23)	38.74 (11.69)	10.94***	PC < AC < GC
French word reading accuracy	23.81 (18.21)	30.91 (17.10)	33.45 (11.24)	7.29***	PC < AC = GC
Grade 3					
Age (in months)	105.07 (4.37)	104.65 (3.58)	105.61 (3.80)	0.35	
English vocabulary	141.67 (19.84)	150.29 (15.00)	154.05 (10.00)	3.65*	PC < AC < GC
French vocabulary	52.33 (19.96)	72.29 (25.07)	82.23 (23.86)	9.24***	PC < AC = GC
English word reading accuracy	47.95 (8.65)	54.37 (7.75)	57.32 (5.30)	9.08***	PC < AC < GC
French word reading accuracy	65.97 (27.63)	66.83 (20.81)	73.36 (20.99)	0.70	
English word reading fluency	56.57 (13.05)	68.21 (8.21)	68.50 (10.60)	8.75***	PC < AC < GC
French word reading fluency	50.76 (12.40)	60.92 (10.30)	62.00 (8.65)	7.75***	PC < AC < GC
English reading comprehension	20.52 (10.35)	30.21 (8.86)	38.77 (6.70)	23.45***	PC < AC < GC
French reading comprehension	11.10 (3.92)	18.37 (2.53)	28.82 (4.88)	114.77***	PC < AC < GC

[a]Equal sign indicates non-significant difference, and less-than symbol indicates $p < 0.05$ or less. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

French reading and language measures in each grade. Univariate analyses were computed for tasks tested at one time point only (i.e., English and French phonological awareness, English and French word reading fluency, and English and French reading comprehension). **Table 5** shows the mean raw scores and standard deviations of the English and French reading and

TABLE 4 | The overlap and distribution of poor reading comprehension in English and French.

Comprehender subgroup	Poor French	Average French	Good French	Not selected for analysis	Total
	n (%)	n (%)	n (%)	n (%)	
Poor English	10 (41.7%)	2 (8.3%)	1 (4.2%)	11 (45.8%)	24
Average English	2 (8.3%)	0 (0%)	1 (4.2%)	21 (87.5%)	24
Good English	2 (8.3%)	3 (12.5%)	14 (58.3%)	5 (20.8%)	24
Not selected for analysis	10 (9.3%)	19 (17.6%)	8 (7.4%)	71 (65.7%)	108
Total	24	24	24	108	180

χ^2 (1, N = 180) = 14.02, p < 0.001.

TABLE 5 | Means (standard deviations) and comparisons of poor comprehenders in English and French, poor comprehenders in English only, poor comprehenders in French only, and controls on English and French measures in grade 1 and grade 3.

Measure	PCB	PCE	PCF	Control group		
	M (SD)	M (SD)	M (SD)		F	Pairwise comparisons
Grade 1						
English phonological awareness	10.00 (7.20)	14.00 (5.40)	13.70 (4.03)	18.20 (8.72)	2.62	
French phonological awareness	11.80 (6.70)	10.64 (5.45)	7.90 (3.60)	11.30 (6.04)	0.98	
English vocabulary	97.80 (19.80)	112.27 (19.04)	116.20 (19.64)	137.00 (13.57)	9.05***	PCB, PCE < C
French vocabulary	26.20 (11.19)	30.27 (11.44)	28.80 (11.24)	40.70 (20.40)	2.15	
English word reading accuracy	28.00 (8.40)	29.60 (10.44)	26.50 (11.41)	31.30 (11.01)	0.39	
French word reading accuracy	28.33 (19.47)	35.55 (17.77)	22.40 (19.29)	31.90 (13.25)	1.05	
Grade 3						
English vocabulary	120.00 (18.08)	127.73 (20.69)	145.80 (18.84)	158.50 (7.99)	13.47***	PCB < PCF, C; PCE < C
French vocabulary	47.20 (22.09)	69.91 (29.40)	48.20 (21.92)	73.10 (24.62)	3.04*	PCB < PCE, C; PCF < C
English word reading accuracy	47.22 (8.69)	51.90 (7.70)	46.40 (10.26)	55.50 (5.52)	2.64	
French word reading accuracy	73.30 (28.40)	74.25 (25.98)	60.67 (19.26)	65.90 (21.03)	0.89	
English word reading fluency	54.70 (20.82)	59.18 (12.98)	48.90 (11.21)	67.10 (11.08)	2.79	
French word reading fluency	50.40 (19.34)	54.36 (12.73)	45.20 (12.59)	58.80 (8.44)	0.17	
English reading comprehension	14.40 (2.59)	16.91 (2.21)	20.00 (11.17)	41.00 (3.53)	38.83***	PCB, PCE, PCF < C; PCB < PCF
French reading comprehension	11.90 (3.07)	15.82 (4.24)	10.50 (4.70)	28.00 (3.74)	37.84***	PCB, PCF < PCE, C; PCE < C

PCB, poor comprehenders in both English and French; PCE, poor comprehenders in English only; PCF, poor comprehenders in French only; C, control group. *p < 0.05; ***p < 0.001.

language measures for each group in grade 1 and grade 3, as well as comparisons across groups.

As expected, there were no significant differences between the four groups on the word reading measures used to select comprehender groups, word reading accuracy and fluency, for both English and French in grade 3, and consistent findings were revealed retrospectively for English and French word reading accuracy in grade 1. Similarly, the groups did not differ significantly on English and French phonological awareness in grades 1 and 3.

Results of univariate analyses showed that there was a significant overall group effect for English reading comprehension, $F(3,41) = 38.83$, $p < 0.001$, $\eta_p^2 = 0.76$ and French reading comprehension, $F(3,41) = 37.84$, $p < 0.001$, $\eta_p^2 = 0.76$. Tukey's HSD *post hoc* comparisons showed that the PCB, PCE, and PCF groups performed worse than the control group on English reading comprehension in grade 3. The PCB group also scored significantly lower than the PCF group on English reading comprehension. For French reading

comprehension in grade 3, all three poor comprehender groups (PCB, PCE, and PCF) scored significantly lower than the control group, with the PCF group also scoring lower than PCE group.

There was a significant overall group effect for English vocabulary, Wilks' $\Lambda = 0.41$, $F(6,70) = 6.64$, $p < 0.001$, $\eta_p^2 = 0.36$, and French vocabulary, Wilks' $\Lambda = 0.29$, $F(6,72) = 3.60$, $p < 0.05$, $\eta_p^2 = 0.22$. Univariate tests revealed that the four groups differed significantly in English vocabulary in grade 1, $F(3,41) = 9.05$, $p < 0.001$, $\eta_p^2 = 0.43$, and in grade 3, $F(3,41) = 13.47$, $p < 0.001$, $\eta_p^2 = 0.54$. Tukey's HSD *post hoc* comparisons showed that children in the PCB and PCE groups scored significantly lower than the control group on English vocabulary in grades 1 and 3. However, in grade 3, the PCB group also scored lower than the PCF group on English vocabulary. The univariate tests for French vocabulary found no significant difference between groups on grade 1 French vocabulary, but there were significant group differences on French vocabulary in grade 3, $F(3,41) = 3.04$, $p < 0.05$, $\eta_p^2 = 0.20$. The *post hoc* test for French vocabulary showed that the PCB and PCF groups had significantly lower

scores than control groups on French vocabulary in grade 3. The PCB group also had lower French vocabulary scores than the PCE group in grade 3.[2]

DISCUSSION

The aim of the present study was to investigate correlates and overlap of reading comprehension difficulties for bilingual poor comprehenders who are exposed to English, the societal language, and French, the language of classroom instruction. By identifying poor comprehenders of both English and French, we were able to determine to what extent poor comprehenders in English, a primary language, are also poor comprehenders in French, an additional language.

We found that there is a moderate degree of overlap in comprehension difficulties in English and French among poor comprehenders with equivalent amounts of exposure to French, with a prevalence rate of 41.7% in our sample. However, our findings also indicate that children who have reading comprehension difficulties in one language do not necessarily have difficulties in another. In addition, we found that English and French vocabulary was a strong and persistent indicator of reading comprehension difficulties in the same language for poor comprehenders of English, French, and both English and French.

Consistent with previous studies, results demonstrate that deficits in oral language are characteristic of children with poor reading comprehension (e.g., Nation et al., 2004, 2010; Catts et al., 2006). Building on previous work (D'Angelo et al., 2014), we found that poor comprehenders of English who received classroom instruction in French demonstrated concurrent vocabulary weaknesses in English and French relative to average and good comprehenders, despite comparable word decoding skills. Lower English vocabulary scores distinguished poor comprehenders from average and good comprehenders, whereas lower French vocabulary scores distinguished poor comprehenders from good comprehenders but not from average comprehenders. Similarly, for children identified in French, poor comprehenders differed from average and good comprehenders on English vocabulary, and from good comprehenders, but not average comprehenders on French vocabulary. These findings suggest that the average comprehenders in this study may have not yet reached a level of French proficiency needed to move beyond the performance of the poor comprehenders on French vocabulary. Vocabulary acquisition in French, an additional language, may be more challenging for immersion children because of their limited exposure to French outside of the classroom. Future research should include measures of cognitive abilities, such as phonological short-term memory that may be better at distinguishing group differences in the early grades (Farnia and Geva, 2011).

Regardless of English or French identification, the retrospective analyses indicated that differences between the three comprehender groups in English and French vocabulary

were apparent in grades 1 and 3, with no group differences on English and French phonological awareness in grade 1. These findings clearly demonstrate that poor comprehenders' oral language weaknesses are evident in the early stages of learning to read in both English and French. Although our study examines poor comprehenders in a bilingual context, these results are strikingly similar to findings reported by Catts et al. (2006) and Nation et al. (2010) and confirm that vocabulary weaknesses are apparent before poor comprehenders' reading comprehension difficulties emerge. However, our study also found that there were differences between poor and average and good comprehenders identified in French on word reading measures in grade 1 and grade 3, indicating that different skills may lead to poor reading comprehension in English and French, and French reading comprehension may be more dependent on word level skills.

This study is the first to demonstrate that children with poor reading comprehension may experience difficulties with comprehension in English, in French, or in both English and French. Of these groups, children who were poor comprehenders in both English and French consistently scored the lowest on English vocabulary in grade 1 and grade 3 and in French vocabulary in grade 3 suggesting that severe English vocabulary weaknesses in poor comprehenders may contribute to comprehension difficulties in English and French. While there were no significant group differences found on phonological awareness, word reading and word fluency tasks, it is interesting to note that the poor comprehenders of both English and French, who were the poorest on English and French reading comprehension, also scored the lowest on all English and French reading and language measures in both grades 1 and 3. Results provide support for the linguistic interdependence hypothesis and suggest that children with poor reading comprehension in L1 may be at risk for being a poor comprehender in L2.

We found that 41.7% of children classified as poor comprehenders in grade 3 were poor comprehenders of both English and French. As expected, this overlap is less than reported in previous studies (e.g., Tong et al., 2017) in part due to differences in the approach to defining poor comprehender groups. More specifically, whereas most previous studies have defined poor comprehender groups based on a cut-off score on word reading, reading comprehension, or both, the present study utilized a regression method to identify poor comprehenders based on the relative discrepancy between wording reading, word reading fluency, and reading comprehension, while controlling for age and non-verbal reasoning, therefore, avoiding overidentification and narrowing the sample of children who qualify for poor comprehender status.

However, it could be argued that the overlap between English and French poor comprehender status should be greater given that English and French are alphabetic orthographies and share many linguistic features. It is worth noting that children in this study had been receiving classroom instruction in French for approximately 3 years at the time of comprehender classification. It is possible that children's poor comprehension in French would have been more apparent had they been exposed to French for a longer period of time. This explanation is consistent with that of previous research, which has demonstrated that

---

[2]Due to the small group sizes, equivalent non-parametric tests were calculated for each analysis. The Kruskal–Wallis test, used for comparing two or more independent samples, confirmed our parametric results.

relative to poor decoders, poor comprehenders' difficulties with reading comprehension emerge around the age 10, when performance in reading comprehension is equally accounted for by oral language and decoding skills (e.g., Elwér et al., 2013). Therefore, it seems plausible that there would be a greater overlap of poor comprehender status with more exposure to the French language in spoken and written form. Further research is needed to investigate the overlap of English and French reading comprehension difficulties in the later elementary grades, as decoding becomes more automatized and greater variance is accounted for by oral language skills.

The current study examined the learning needs of poor comprehenders in immersion education and has important implications for the assessment and remediation of reading comprehension difficulties in emerging bilingual learners. Our findings demonstrate that poor comprehenders exhibit pervasive oral language difficulties from the onset of reading that manifest similarly in English, their primary language, and French, the language of instruction. Furthermore, the results suggest that it is possible for children to experience poor reading comprehension in one language but be relatively good at comprehension in another language. Since many children begin French immersion with limited levels of French language proficiency, it is beneficial to gather information on children's reading and language abilities with parallel measures in English and French. Limiting assessment to French, an additional language, may underestimate children's reading and language ability or misattribute reading difficulties to a lack of French proficiency (Geva and Herbert, 2012).

This research also suggests that intervention strategies should be targeted at poor comprehenders' underlying language difficulties regardless of language of instruction. While there have been relatively few intervention studies with poor comprehenders, existing studies have shown that intervention practices that promote oral language skills and text comprehension strategies are effective supports for monolingual children with poor reading comprehension (Snowling and Hulme, 2012). Evidently, there is a need for future intervention research that fosters the development of children's oral language skills in immersion programs.

There are some limitations of the current study that should be noted. First, the sample of poor comprehenders identified within the three subgroups (i.e., PCB, PCE, PCF) was small, which limits the generalizability of our findings. However, obtaining a large sample of poor comprehenders is particularly challenging in a bilingual educational context. Our study is among the few longitudinal studies that have examined bilingual poor comprehenders' reading and language skills in both languages over time. Given the attrition of students in French immersion (e.g., Chen et al., 2019) and the prevalence rate of poor comprehenders in middle elementary years at approximately 10% (e.g., Nation and Snowling, 1998; Clarke et al., 2010), our sample size may be considered representative of poor comprehenders in a bilingual context. Nevertheless, larger sample sizes for the subgroups of poor comprehenders would benefit future work.

Reading comprehension is a complex process that involves the coordination of various skills that are assessed differently across measures of reading comprehension. In the present study, we used a single standardized measure of reading comprehension. Although the use of this standardized test makes our sample of poor comprehenders comparable to those in the existing monolingual literature (e.g., Tong et al., 2014), results reported in this study need to be replicated with more varied reading comprehension measures to disentangle whether poor comprehenders score low on reading comprehension because they do not understand the text or because they are unable to read the question. Similarly, the use of a single measure of vocabulary knowledge may not fully capture the influence of other language skills on reading comprehension, such as vocabulary depth, listening comprehension, morphological awareness, and inference (Nation and Cocksey, 2009; D'Angelo and Chen, 2017).

Another limitation is that approximately 25% of the children identified as poor comprehenders in either English, French, or both were exposed to another language at home in addition to English. While this sample is representative of students enrolled in French immersion programs in Canada, there is a need for further research to explore whether significant differences exist between children identified as poor comprehenders from English monolingual backgrounds and those who speak additional languages.

Finally, there is some difficulty in interpreting poor comprehender status in French only, particularly for children in this study who grew up in an English-speaking community. Poor reading comprehension in French may not be attributed to a language impairment or limited proficiency in French but associated with children's lack of motivation to learn in an L2. Evidently, there is a need for further research to explore the role of motivation in L1 and L2 reading comprehension for children enrolled in immersion programs.

Taken together, the present study demonstrates that poor comprehenders experience similar and persistent difficulties with components of language in both English, a primary language, and French, an additional language, that are present in the early stages of reading development, and therefore, likely indicators of later reading comprehension difficulties in both languages. These results also show while there is a moderate degree of overlap in English and French reading comprehension difficulties, not all poor comprehenders of English are poor comprehenders of French, suggesting that somewhat different skills may be involved in comprehending text in English and French.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Toronto Research Ethics Board.

Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

ND'A and XC contributed to the conception and design of the study. ND'A and KK organized data collection and managed the database and performed the statistical analyses. ND'A wrote the first draft of the manuscript. KK and XC wrote sections of the manuscript. All authors contributed to manuscript revisions and read and approved the submitted version.

REFERENCES

August, D., and Shanahan, T. (2006). *Developing Literacy in Second-Language Learners: A Report of the National Literacy Panel on Minority-Language Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.

Au-Yeung, K., Hipfner-Boucher, K., Chen, X., Pasquarella, A., D'Angelo, N., and Deacon, S. H. (2015). Development of English and French language and literacy skills in EL1 and EL French immersion students in the early grades. *Read. Res. Q.* 50, 233–254. doi: 10.1002/rrq.95

Catts, H. W., Adlof, S. M., and Weismer, S. E. (2006). Language deficits in poor comprehenders: a case for the simple view of reading. *J. Speech Lang. Hear. Res.* 49, 278–293. doi: 10.1044/1092-4388(2006/023)

Catts, H. W., Compton, D., Tomblin, J. B., and Bridges, M. (2012). Prevalence and nature of late emerging reading disabilities. *J. Educ. Psychol.* 104, 166–181. doi: 10.1037/a0025323

Chen, X., Burchell, D., and Sinay, E. (2019). Demographic characteristics, engagement, and achievement of Canadian students in French immersion and extended dual-language programs. *Paper Presented at the 2019 Annual Meeting of the American Educational Research Association*, Toronto, ON.

Clarke, P., Snowling, M., Truelove, E., and Hulme, C. (2010). Ameliorating children's reading comprehension difficulties: a randomized controlled trial. *Psychol. Sci.* 21, 1106–1116. doi: 10.1177/0956797610375449

Comeau, L., Cormier, P., Grandmaison, E., and Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *J. Educ. Psychol.* 91, 29–43. doi: 10.1037/0022-0663.91.1.29

Cummins, J. (1984). *Bilingualism and Special Education: Issues in Assessment and Pedagogy*. Clevedon: Multilingual Matters.

Da Fontoura, H. A., and Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese-English Canadian children. *Read. Writ. Interdiscip. J.* 7, 139–153. doi: 10.1007/bf01026951

D'Angelo, N., and Chen, X. (2017). Language profiles of poor comprehenders in English and French. *J. Res. Read.* 40, 153–168. doi: 10.1111/1467-9817.12084

D'Angelo, N., Hipfner-Boucher, K., and Chen, X. (2014). Poor comprehenders in French immersion: implications for identification and instruction. *Perspect. Lang. Lit.* 40, 32–37.

D'Angelo, N., Hipfner-Boucher, K., and Chen, X. (2017). Predicting growth in English and French vocabulary: the facilitating effects of morphological and cognate awareness. *Dev. Psychol.* 53, 1242–1255. doi: 10.1037/dev0000326

Deacon, S. H., and Kirby, J. R. (2004). Morphological: is it more than phonological? *Appl. Psycholinguist.* 25, 223–238. doi: 10.1017/s0142716404001110

Dunn, L. M., and Dunn, D. M. (2007). *Peabody Picture Vocabulary Test*, 4th Edn. Bloomington, MN: NCS, Pearson.

Dunn, L. M., Theriault-Whalen, C. M., and Dunn, L. M. (1993). *Échelle de Vocabulaire en Images Peabody (EVIP)*. Richmond Hill, ON: Psycan.

Elwér, Å., Keenan, J. M., Olson, R. K., Byrne, B., and Samuelsson, S. (2013). Longitudinal stability and predictors of poor oral comprehenders and poor decoders. *J. Exp. Child Psychol.* 115, 497–516. doi: 10.1016/j.jecp.2012.12.001

Farnia, F., and Geva, E. (2011). Cognitive correlates of vocabulary growth in English language learners. *Appl. Psycholinguist.* 32, 711–738. doi: 10.1017/s0142716411000038

Geva, E., and Clifton, S. (1994). The development of first and second language reading skills in early French immersion. *Can. Mod. Lang. Rev.* 50, 646–667.

FUNDING

ACKNOWLEDGMENTS

Geva, E., and Herbert, K. (2012). "Assessment and Interventions in English Language Learners with LD," in *Learning about Learning Disabilities*, 4th Edn, eds Bernice, and Deborah, (San Diego, CA: Academic Press), 271–298.

Gottardo, A., and Mueller, J. (2009). Are first and second language factors related in predicting L2 RC? *J. Educ. Psychol.* 101, 330–344. doi: 10.1037/a0014320

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104

Koda, K. (2008). "Contributions of prior literacy experience in learning to read in a second language," in *Learning to Read Across Languages: Cross-Linguistic Relationships in First and Second-Language Literacy Development*, eds K. Koda, and A. M. Zehler (New York, NY: Routledge).

LeBlanc, R., and Seguin, H. (1996). "Les congeneres homographes et parographes anglais-francais," in *Twenty-Five Years of Second Language Teaching at the University of Ottawa*, (Ottawa, CA: University of Ottawa Press), 69–91.

Lesaux, N. K., Lipka, O., and Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Read. Writ. Interdiscip. J.* 19, 99–131. doi: 10.1007/s11145-005-4713-6

Li, M., and Kirby, J. R. (2014). Unexpected poor comprehenders among adolescent ESL students. *Sci. Stud. Read.* 18, 75–93. doi: 10.1080/10888438.2013.775130

MacCoubrey, S. J., Wade-Woolley, L., Klinger, D., and Kirby, J. R. (2004). Early identification of at-risk L2 readers. *Can. Mod. Lang. Rev.* 61, 11–28. doi: 10.3138/cmlr.61.1.11

MacGinitie, W. H., and MacGinitie, R. K. (1992). *Gates-MacGinitie Reading Tests*, 2nd Edn. Toronto, CA: Nelson Canada.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., and Dreyer, L. G. (2000). *Gates-MacGinitie Reading Tests*, 4th Edn. Itasca, IL: Riverside Publishing.

Mancilla-Martinez, J., and Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: the case of Spanish-speaking language minority learners. *J. Educ. Psychol.* 102, 701–711. doi: 10.1037/a0019135

Manis, F. R., and Lindsey, K. A. (2010). "Cognitive and oral language contributors to reading disabilities in Spanish-English bilinguals," in *Language and Literacy Development in Bilingual Settings*, eds A. Y. Durgunoglu, and C. Goldenberg (New York, NY: Guilford press).

McBride-Chang, C., Shu, H., Chan, W., Wong, T., Wong, A. M. Y., Zhang, Y., et al. (2013). Poor readers of Chinese and English: overlap, stability, and longitudinal correlates. *Sci. Stud. Read.* 17, 57–70. doi: 10.1080/10888438.2012.689787

Naglieri, J. A. (1985). *Matrix Analogies Test – Short Form*. San Antonio, TX: The Psychological Corporation.

Nation, K., Clarke, P., Marshall, C. M., and Durand, M. (2004). Hidden language impairments in children: parallels between poor reading comprehension and specific language impairments? *J. Speech Lang. Hear. Res.* 47, 199–211. doi: 10.1044/1092-4388(2004/017)

Nation, K., and Cocksey, J. (2009). The relationship between knowing a word and reading it aloud in children's word reading development. *J. Exp. Child Psychol.* 103, 296–308. doi: 10.1016/j.jecp.2009.03.004

Nation, K., Cocksey, J., Taylor, J. S., and Bishop, D. V. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *J. Child Psychol. Psychiatr.* 51, 1031–1039. doi: 10.1111/j.1469-7610.2010.02254.x

Nation, K., and Snowling, M. J. (1998). Individual differences in contextual facilitation: evidence from dyslexia and poor reading comprehension. *Child Dev.* 69, 996–1011. doi: 10.1111/j.1467-8624.1998.tb06 157.x

Paradis, J., Genesee, F., and Crago, M. (2010). *Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning*, 2nd Edn. Baltimore, MD: Brookes Publishing.

Roy, C., and Labelle, M. (2007). Connaissance de la morphologie dérivationnelle chez les francophones et non-francophones de 6 à 8 ans. *Can. J. Appl. Linguist.* 10, 263–291.

Seymour, P. H. K., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859

Shum, K.-K., Ho, C., Siegel, L., and Au, T.-K.-F. (2016). First-language longitudinal predictors of second-language literacy in young L2 learners. *Read. Res. Q.* 51, 323–344. doi: 10.1002/rrq.139

Snowling, M. J. (2000). *Dyslexia*, 2nd Edn. Oxford: Blackwell.

Snowling, M. J., and Hulme, C. (2012). The nature and classification of reading disorders: a commentary on proposals for DSM-5. *J. Child Psychol. Psychiatr.* 53, 593–607. doi: 10.1111/j.1469-7610.2011.02495.x

Tong, X., Deacon, S. H., and Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: another piece of the puzzle. *J. Learn. Disabil.* 47, 22–33. doi: 10.1177/0022219413509971

Tong, X., Deacon, S. H., Kirby, J. R., Cain, K., and Parrila, R. (2011). Morphological awareness: a key to understanding poor reading comprehension in English. *J. Educ. Psychol.* 103, 523–534. doi: 10.1037/a002 3495

Tong, X., McBride, C., Shu, H., and Ho, C. (2017). Reading comprehension difficulties in Chinese-English bilingual children. *Dyslexia* 24, 59–83. doi: 10. 1002/dys.1566

Tong, X., Tong, X., and McBride-Chang, C. (2015). A tale of two writing systems: double dissociation and metalinguistic transfer between Chinese and English word reading among Hong Kong children. *J. Learn. Disabil.* 48, 130–145. doi: 10.1177/0022219413492854

Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (1999). *Test of Word Reading Efficiency*. Austin, TX: PRO-ED Publishing, Inc.

Wagner, R. K., Torgesen, J. K., and Rashotte, C. A. (1999). *Comprehensive test of Phonological Processing*. Austin, TX: PRO-ED.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., and Pearson, N. A. (2013). *CTOPP-2: Comprehensive test of Phonological Processing*, 2nd Edn. Austin, TX: PRO-ED.

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Rolling Meadow, IL: Riverside Publishing.

Reading in English as a Foreign Language by Spanish Children With Dyslexia

Paz Suárez-Coalla*, Cristina Martínez-García and Andrés Carnota

Department of Psychology, University of Oviedo, Oviedo, Spain

It has been reported that children with dyslexia have difficulties with learning a second language. The English alphabetic code is opaque, and it has been stated that deep orthographies cause important problems in children with dyslexia. Considering the strong differences between the Spanish and English orthographic systems, we predicted English reading problems in Spanish-speaking children with dyslexia. The current study focused on English as a foreign language in a group of 22 Spanish children with dyslexia (8–12 year olds), compared to a control group matched for age, gender, grade, and socioeconomic status. The objective was to identify the main difficulties that Spanish-speaking children with dyslexia demonstrate during English reading, to develop specific teaching programs. Participants were given four tasks related to reading: discrimination of phonemes, visual lexical decision, reading aloud, and oral vs. written semantic classification. The results suggest that children with dyslexia demonstrate problems in using English grapheme–phoneme rules, forcing them to employ a lexical strategy to read English words. However, they also showed difficulties in developing orthographic representations of words. Finally, they also exhibited problems with oral language, demonstrating difficulties accessing semantic information from an auditory presentation.

Keywords: English as a foreign language, dyslexia, reading, Spanish, children

INTRODUCTION

Developmental Dyslexia

Developmental dyslexia is considered a neurobiological condition characterized by specific and pronounced difficulty in reading and writing acquisition. This condition results in persistent accuracy and speed deficits in both reading and writing competencies (Grainger et al., 2003; Lyon et al., 2003; Suárez-Coalla and Cuetos, 2012, 2015; Afonso et al., 2019). The origin of literacy acquisition problems has been repeatedly attributed to deficits in phonological processing or the ability to identify and manipulate speech sounds (Goswami and Bryant, 1990; Stanovich and Siegel, 1994; Serrano and Defior, 2008). Moreover, recent studies suggest that the phonological deficit could be partially caused by certain subtle disorders in sound perception, preventing children with dyslexia from developing good phonological representation (Goswami, 2002; Boets et al., 2006; Beattie and Manis, 2012; Cuetos et al., 2018). Consequently, disorders in sound perception could

determine phonological awareness and literacy acquisition, which could cause more pronounced and profound problems for foreign language (FL) learning.

Reading Acquisition

To achieve reading competence in the first/native language (L1), it is necessary to acquire the grapheme–phoneme (G–P) conversion rules, in addition to developing orthographic representations of intermediate units (groups of graphemes) and words. The development of orthographic representations of words is particularly pertinent in deep orthographic systems. It is well-recognized that the first step in learning to read is to learn the alphabetic code (Ehri, 1987, 1998; Share, 1995). This knowledge of G–P correspondences is critical, as it permits people to decode the written language. However, this knowledge is not sufficient in and of itself to constitute fluent reading skills. To develop reading fluency, we also need to store orthographic representations of words. This facilitates direct, smooth, and fast reading, without having to convert each grapheme into its corresponding phoneme. According to the self-teaching hypothesis, and supported by many studies, the accurate and repeated decoding of words facilitates the storing of the orthographic representation of those words in memory (Ehri and Roberts, 1979; Reitsma, 1989; Share, 1999; Cunningham, 2006; Kyte and Johnson, 2006; Maloney et al., 2009). In this sense, the acquisition and automation of the alphabetic code is crucial to obtaining an appropriate and robust orthographic representation of new words. However, the characteristics of the orthographic system seem to determine the evolution of the reading strategies. In deep orthographic systems, like English, G–P irregularities seem to force (from an early age) the development of orthographic representations of intermediate units (i.e., rhymes, syllables, and morphemes) and words (Wang et al., 2012). By contrast, in transparent orthographic systems, like Spanish or Italian, the G–P correspondence rules are very easy to learn, and children decode them accurately from the outset (Seymour et al., 2003; Cuetos and Suárez-Coalla, 2009). Nevertheless, even in transparent orthographies, children also develop representations for intermediate units (Burani et al., 2002; Cuetos and Suárez-Coalla, 2009), and for whole words (Suárez-Coalla et al., 2016). It has been reported that several variables modulate orthographic storing, including, for example, syllable structure, context-dependent graphemes, and phonological or semantic knowledge of new words (Bowey, 1995; Walley et al., 2003; Ricketts et al., 2007; Wang et al., 2013; Álvarez-Cañizo et al., 2018). The development of reading fluency therefore depends on many variables.

Regarding dyslexia, studies about reading difficulties support the idea that children with dyslexia demonstrate problems in learning the alphabetic code and creating orthographic representations of words. Evidently, if they do not successfully learn the alphabetic code – because of difficulties in phonological processing – and make mistakes when reading words, they will have problems storing correct representations (Bruck, 1992; Vellutino et al., 2004). It has been reported that children with dyslexia remain inaccurate and slow after a significant number of decoding opportunities (Hogaboam and Perfetti, 1978; Reitsma, 1983; Manis, 1985; Ehri and Saltmarsh, 1995; Cao et al., 2006;

Martens and de Jong, 2008; Clements-Stephens et al., 2012; Suárez-Coalla et al., 2014b). These difficulties seem to lead to the use of a sublexical reading strategy. Children without reading difficulties, by contrast, use lexical reading strategies from an early age (Suárez-Coalla and Cuetos, 2012; Davies et al., 2013).

Furthermore, cross-linguistic studies have reported that the reading performance of people with dyslexia varies depending on the orthographic system, resulting in diverse behavioral manifestations, in spite of dyslexia's common neurological origin. In particular, as a consequence of orthographic depth, it has been noted that dyslexic reading accuracy problems are more pronounced in deep orthographies (e.g., English) than in shallow ones (e.g., Spanish or Italian) (Wimmer, 1993; Wimmer and Goswami, 1994; Ziegler and Goswami, 2005).

Reading slowness constitutes the main marker of dyslexia in shallow orthographies (De Jong and van der Leij, 2002; Constantinidou and Stainthorp, 2009; Suárez-Coalla and Cuetos, 2012). In Spanish, slowness seems to be a consequence of using a sublexical strategy for reading aloud and the lack of mastery of the G–P rules. Specifically, Spanish-speaking children with dyslexia show a significant effect based on the length of the stimuli (a marker of a sublexical strategy) and continue to show the length effect after repeated exposure to words (Suárez-Coalla et al., 2014a,b; Martínez-García et al., 2019). It is thought that this persistent length effect is an indicator of the absence of orthographic information (Suárez-Coalla et al., 2014a,b; Martínez-García et al., 2019).

Reading in English as a Foreign Language

Children all over the world learn English as an FL (EFL) at an early age (Bonifacci et al., 2017). Schools try to prepare children for a global society, and the language barrier constitutes a challenge to children with dyslexia. They must learn two different – sometimes widely divergent – alphabetic codes (e.g., English vs. Spanish). Spanish has a highly transparent orthography, with high correspondence between graphemes and phonemes (Seymour et al., 2003). Spanish speakers pronounce the majority of graphemes without variation, except for three consonants (c, g, and r). There are certain rules, however, which regulate the pronunciation of these consonants in relation to accompanying vowels and their position in a word. Moreover, Spanish has five double-letter graphemes (ll, rr, ch, gu, and qu), which only appear at the beginning of the syllable. In addition, there are only five vowels (a, e, i, o, and u), and their pronunciation does not vary. In general, the method of reading instruction in Spanish is phonetic–syllabic: children learn single letters and their sounds and then combine letters to read syllables and words. Therefore, because of the consistency of Spanish orthography, children without problems achieve reading accuracy very early on (i.e., during the first year of exposure to reading) (Seymour et al., 2003). However, this is not the case for children with dyslexia (Davies et al., 2007; Suárez-Coalla and Cuetos, 2012).

By contrast, English orthography is more irregular. In the English alphabet, there are 26 letters (21 consonants and 5 vowels, 6 if we consider "y," also a vowel when it is the only "vowel"

in a word, e.g., "sky"), but there are more than 40 consonant and vowel sounds. In some words (e.g., "best"), the number of letters and sounds is the same (four letters and four sounds). In other words, however (e.g., "green"), the number of letters and sounds is different (five letters and four sounds). In addition, some words have the same pronunciation but different spellings (e.g., "know" vs. "no"), and some have the same spellings but different pronunciation (e.g., "read:" infinitive vs. past tense) (Marks, 2007). These irregularities make it difficult to read EFL, particularly for children with dyslexia.

In Spain, children begin to be informally exposed to English from the beginning of preschool, when they are around 3 years old. However, English is introduced in a more formal and academically rigorous way in Year 1 of Primary School, at the age of 6. Currently, children receive EFL lessons for approximately 4 h/week. In addition, increasing amounts of bilingual education are being introduced in Spain, with ∼50% of subjects being taught in the English language. To teach reading in English, instructors mainly use a global method – introducing meaning, pronunciation, and spelling at the same time. This constitutes a significant challenge.

It is well-known that English reading causes particular difficulties for children with dyslexia (Nijakowska, 2010), and in Spain, these difficulties are often noted by parents and teachers. However, these difficulties are rarely assessed by clinicians and speech therapists, probably due to the absence of formal EFL testing, as well as the absence of specific training in EFL testing, and the traditional priority given to L1, as mentioned by Helland and Kaasa (2005) in the Norwegian context. Therefore, research in this field is critically necessary to develop an understanding of how Spanish children with dyslexia tackle reading in EFL.

Pioneering and influential studies about FL learning difficulties have advanced the Linguistic Coding Deficits Hypothesis (LCDH) (Sparks et al., 1999, 2012). This asserts that FL acquisition is related to phonological, orthographic, syntactic, and semantic skills in L1. The LCDH suggests that FL learning is built on L1 skills. Therefore, the strength of the L1 codes determines the student's future success in FL learning. The assumptions derived from the LCDH have attracted the attention of multiple researchers. Specifically, it has been argued that people with reading problems in L1 will be prone to reading problems in an FL – that is, early problems with phonological and orthographic processes in L1 will be transferred to the FL (Chodkiewicz, 1986; Durgunoglu et al., 1993; Cisero and Royer, 1995; Da Fontoura and Siegel, 1995; Geva et al., 1997; Comeau et al., 1999; Dufva and Voeten, 1999; August et al., 2001; Kahn-Horwitz et al., 2006).

Accordingly, studies addressing reading in EFL (China, Italy, Norway, and Poland, etc.) have reported worse English reading performance in people with dyslexia than in typical readers, regardless of the characteristics of the L1 orthographic system (Chinese: Ho and Fong, 2005; Chung and Ho, 2010; Hebrew: Oren and Breznitz, 2005; Italian: Palladino et al., 2013; Norwegian: Helland and Kaasa, 2005; Polish: Lockiewicz and Jaskulska, 2016).

Lockiewicz and Jaskulska (2016) performed a study with Polish adolescents with dyslexia (aged 16–18), in which they were asked to read English words and pseudo-words. Significant differences were found between typical readers and adolescents with dyslexia. Students with dyslexia showed less accuracy and a slower reading speed than the control group, in both English words and pseudo-words. In another study, Palladino et al. (2013) compared Italian-speaking children with and without dyslexia (aged 12–14). In this study, children were also asked to read English words and pseudo-words. The Italian children with dyslexia showed poor reading of English words; however, contrary to the results reported by Lockiewicz and Jaskulska (2016), they seemed to manage English G–P rules because they showed a high level of accuracy when reading pseudo-words. Considering Norwegian students (aged 12), Helland and Kaasa (2005) found significant differences between groups in literacy tasks, who were tested on spelling, translation, and reading skills. In the context of the Chinese language, where the script differs significantly from that of English, Ho and Fong (2005) found that primary school children with dyslexia performed significantly worse than the control group in several English measures (vocabulary, reading, and phonological processing tasks). In addition, they found that phonological skills correlated with English reading.

Primarily, results have suggested that reading problems in L1 are a predictor of reading difficulties in EFL, probably due to a common cause. However, Miller-Guron and Lundberg (2000) reported surprising results. They found that some Swedish adults with dyslexia demonstrated a preference for reading in English, instead of Swedish (L1), even though Swedish orthography is more transparent than that of English. This phenomenon was termed the "dyslexic preference for English reading," and it was believed to be a consequence of different factors, including age, and EFL exposure (mass media, literature, etc.). It is also believed to be related to a preference for larger orthographic segments, due to their inherent challenges with G–P decoding. These results could be modulated by other variables, but they are not generalizable. In this sense, it is interesting to continue researching about reading performance and strategies in EFL, especially in populations with different L1 orthographic systems (and different sociocultural contexts).

Regarding differences between L1 and English orthographic systems, it is reasonable to anticipate certain difficulties when learning two different alphabetic codes. For example, it must be considered that the complex graphemes of English (e.g., ea, ph, and ow – which do not exist in Spanish) could pose a difficulty to Spanish children. This graphemic complexity effect has been found in French children when reading English words (Commissaire, 2012), suggesting that the identification of complex graphemes competes with the identification of simple graphemes. On the other hand, the phonological representation of the words seems to be activated automatically during the visual recognition of words in bilingual individuals (Dijkstra and van Heuven, 2002); the same holds true for the grapheme–phoneme rules of both languages (Goswami et al., 2001; Jared and Kroll, 2001; Van Wijnendaele and Brysbaert, 2002). This finding suggests that differences in the orthographic systems could cause interferences to readers in EFL, especially when L1 is a transparent

orthography (like Spanish). However, the level of activation could depend on the individual's fluency and experience with languages (Jared and Kroll, 2001). It has been proven, however, that university students who are learners of EFL are sensitive to the morphological structure of English words, indicating that they are able to recode the written word into different grain sizes of psycholinguistic units (Casalis et al., 2015). Moreover, the "grain size accommodation" hypothesis has recently suggested that learning to read in consistent and inconsistent orthographies concomitantly is in fact advantageous to readers (Lallier and Carreiras, 2017). Readers in this context seem to increase their use of phonological strategies in opaque orthographies and lexical processing in transparent orthographies.

To our knowledge, there are no studies about reading in English by Spanish-speaking children with dyslexia. Taking into account previous results, as well as the difficulties reported by teachers and parents, it can be expected that dyslexic Spanish children will have problems with reading in EFL.

The aim of this study was to describe specific difficulties of – and reading strategies used by – Spanish children with dyslexia in EFL reading, compared to typical Spanish readers. Specifically, we tried to determine whether Spanish children with dyslexia were able to use some English G–P rules to read unfamiliar words or, alternatively, whether they had difficulties managing English regularities. We also tried to determine if they had developed the orthographic representations of words or, instead, whether there was some kind of phonological and/or cross-linguistic interference (i.e., if they activate the Spanish phonology when reading in EFL). Furthermore, we intend to evaluate whether problems with the discrimination of phonemes were also noticeable in this population, as it has been argued that auditory deficits exist in people with dyslexia, which probably affect phonological representations and therefore English learning. To achieve our goals, four tasks were performed: discrimination of phonemes (same–different), visual lexical decision-making, reading aloud, and oral vs. written semantic categorization. Participants were native Spanish speakers with and without dyslexia, from 8 to 12 years old, who studied EFL as a compulsory subject at school. We assumed that children with dyslexia would show worse performance in all tasks, with more mistakes and longer reaction times (RTs) than children without dyslexia.

In short, we are seeking to address the following issues related to EFL reading:

- Do children with dyslexia have representations of English phonemes and the ability to discriminate among them?
- Do children with dyslexia know the G–P conversion rules of English and use them to read unknown words?
- Do children with dyslexia developed robust orthographic representations of English words?
- Are children with dyslexia able to access semantic information from oral and/or written presentations? Do they demonstrate differences depending on the form of presentation of the stimuli?

METHODOLOGY

Participants

A total of 44 children (24 boys and 20 girls) between 8 and 12 years of age (M = 10 years, 8 months, SD = 0.8) participated in the study. Half of the participants had dyslexia (DYS), and half were typical readers (CON). Both groups were matched by age, gender, grade, and socioeconomic status. All the participants were native Spanish speakers and had no known motor, cognitive, or perceptual disorders.

Participants without dyslexia were recruited from several primary schools in Asturias (Spain). The children with dyslexia were recruited from the Association of Dyslexia and certain Speech Therapy Centers of Asturias (Spain). Children with dyslexia had previously received the diagnosis of dyslexia, had an intelligence quotient (IQ) of 85 or higher (M = 109; SD = 7.58), according to the Wechsler Intelligence Scale for Children (Wechsler, 2001), and had shown a low phonological awareness performance (in a phoneme omission task created by the authors of the study). The average score in the phonological awareness task was M = 6.67 (out of 10), SD = 1.70. The average score for the typical readers was M = 9.40, SD = 0.87.

Before performing the experimental tasks, a reading battery (PROLEC-R, Cuetos et al., 2007, or PROLEC-SE, Ramos and Cuetos, 2005) was administered to all participants, to confirm the diagnosis of reading difficulties. PROLEC-R and PROLEC-SE yield scores (accuracy and total reading times) for words and pseudo-words. The section of words consists of 40 Spanish words (high and low frequency, short, and long words). The pseudo-words section includes 40 pseudo-words, half of which were short and half of which were long. Children were included in the DYS group if both accuracy and reading speed scored 1.5–2 standard deviations below the age mean, according to age norms provided by PROLEC-R or PROLEC-SE. Meanwhile, children were included in the CON group when they had an age-appropriate score in both sections. Means, standard deviations, and p values for scores obtained in reading assessment tests are provided in **Table 1**.

TABLE 1 | Means, standard deviations and p-values for scores obtained in reading assessment tests.

	DYS mean (SD)	CON mean (SD)	p-value
Age (years)	10.9 (0.9)	10.7 (0.8)	p = 0.67
Reading			
Words			
Accuracy (out of 40)	35.83 (1.86)	39.76 (0.59)	p < 0.001
Speed (s)	62.78 (28.40)	30.22 (13.35)	p < 0.001
Pseudowords			
Accuracy (out of 40)	33.50 (2.85)	38.11 (0.81)	p < 0.001
Speed (s)	80.06 (18.84)	42.25 (9.35)	p < 0.001

Materials and Methods

Four tasks were performed: discrimination of phonemes (same–different), visual lexical decision-making, reading aloud, and oral vs. written semantic categorization. Each task lasted ∼5 min.

Discrimination of Phonemes: Same–Different

The relationship between phonological processing and reading acquisition is well-known (Wagner and Torgesen, 1987; Adams, 1990; Scanlon et al., 2000; Snowling, 2000). In line with the definition provided by Catts et al. (1999), phonological processing includes the perception, storage, recovery, and manipulation of language sounds. The ability to manipulate speech units requires an important level of awareness that words are formed by sublexical segments (i.e., discrete sounds that can be recombined). Different tasks (omission and identification of phonemes, rhyming judgments, word segmentation, and discrimination of phonemes) are used to assess phonological awareness and its relation to reading.

In English, there are more vowel phonemes than in Spanish, some of which have subtle differences that are very difficult for Spanish people to discriminate (e.g., the sound/I/as in "sit" vs. the sound/i:/as in "seat"). This could pose a problem to Spanish children with phonological processing deficits, such as children with dyslexia (Elbro, 1996; Snowling, 2000). By contrast, it has also been suggested that children with dyslexia – who have problems acquiring the phonological categories of L1 – retain sensitivity to universal phonetic boundaries, which are lost in typical phonological acquisition (Serniclaes et al., 2004; Soroli et al., 2010). In this sense, this task aimed to ascertain whether children with dyslexia have problems discriminating English vowel phonemes.

A total of 36 English monosyllabic words were selected. From the selected stimuli, 12 pairs, including pairs featuring the same word, were formed (e.g., hot–hot). From the remaining stimuli, 12 pairs were created, including two different words which only differed in one phoneme (e.g., sheep–ship). In addition, four trials were included at the beginning as practice, to familiarize children with the task.

Participants were orally presented with pairs, and they had to decide whether the pairs contained the same word or different words. They were told that they were going to hear two stimuli, and they were asked to decide, as quickly and accurately as possible, whether they were the same or different by pressing the appropriate key. One key had a green sticker placed on top, which was used to indicate that the words were the same, and another key, which a red sticker placed on top, was used to indicate that the words were different.

To obtain the auditory stimuli, words were recorded with a Zoom H4N recorder and Audix Ht2-P Plantronics microphones. Subsequently, the stimuli were edited with Praat software (Boersma and Weenink, 2019).The experimental task was run on an HP Mini laptop, with the DMDX program (Forster and Forster, 2003). The trial started with a warning tone and an asterisk on the screen, followed by the two words. A silent interval of 500 ms was placed between the two stimuli. Timing started from the onset of the second stimulus. The type of response (correct or incorrect) and RTs were recorded as data. Cronbach's alpha was 0.55.

Reading Aloud

According to dual-process models, reading may be conducted through two different processing routes. The sublexical route uses knowledge about the alphabetical code: the G–P rules of the language. Alternatively, the lexical route makes use of the orthographic representations of words to lexically access their phonological representations (Coltheart and Rastle, 1994; Coltheart et al., 2001). When you have to read an unknown word, you do not have a pre-existing orthographic representation of it. Therefore, you have to use the alphabetical code or use some kind of analogy to obtain the correct pronunciation. With this task, we tried to determine if the children with dyslexia were able to read infrequent and unfamiliar words based on the knowledge of certain G–P conversion rules of the English orthographic system. When reading aloud regular words, it is not necessary to know the orthographic representation of the word or the word meanings. This task could therefore inform us about the ability of dyslexic children to manage English G–P rules.

A total of 24 words were selected. Half of them were high-frequency (HF) words (M = 63,198, SD = 86,807) and were considered familiar to children, as they were drawn from their English schoolbook [e.g., "table" ('teɪbəl)]. The other half of the words were low-frequency (LF) words (M = 1,388, SD = 3,240), previously unknown to the children [e.g., "gable" ('geɪbəl)]. The lexical frequency was obtained from the Hyperspace Analog to Language (HAL) frequency norms (Lund and Burgess, 1996). These frequency norms were based on the HAL corpus, which consists of ∼131 million words gathered across 3,000 Usenet newsgroups during February 1995, cited in *The English Lexicon Project* (Balota et al., 2007). These unknown words were orthographic and phonological neighbors to the known ones, since the two words differed only in a consonant phoneme. The vowel phoneme remained the same (in terms of spelling and pronunciation).

From these words, we created two lists of words matched on frequency and pronunciation, so that children received six known and six unknown words. Each word was presented visually (20-point Arial font) to participants for 4,000 ms. They were asked to read the word aloud as quickly and accurately as possible. RTs were considered – that is, the duration between the onset of the target on the screen and the time when the participants started to articulated the word.

The experimental task was run on an HP Mini laptop, and the responses were recorded in.WAV files with the DMDX program (Forster and Forster, 2003). A trial started with a warning tone and an asterisk on the screen, followed by the word to be read. The sound spectrograms of the recorded responses were analyzed using the CheckVocal application (Protopapas, 2007) to extract accuracy and RTs. Mistakes (self-corrections, substitutions, and regressions) and omitted responses were excluded. Cronbach's alpha was 0.88.

Lexical Decision-Making Task

To perform a visual lexical decision-making task, it is necessary to have developed a robust orthographic representation, especially when it comes to irregular words. On the other hand, it has been reported that the L1 phonology is activated even when an individual is reading an FL. In this sense, the visual decision task has been used to ascertain the influence of L1 in FL word recognition (Elston-Güttler et al., 2005). It is relevant when L1 and FL phonemes differ considerably, as they do in Spanish and English. The objective of this task was to ascertain if children with dyslexia had developed orthographic representations of English words or, alternatively, if they were affected by phonological cross-linguistic interferences.

In this task, the participants had to recognize and decide if a visually presented letter sequence constituted a real word. A total of 32 stimuli were selected, manipulating lexical frequency and length. For the short stimuli, the mean length was 3.75 (SD = 0.43, three to four letters), and for the long stimuli, the mean length was 7.55 (SD = 0.96, six to nine letters). Regarding the frequency values of words, the mean for the HF words was 176,051 (SD = 77,138), and for the LF words, the mean was 6,988 (SD = 3,470). The frequency values were obtained from the HAL frequency norms (Lund and Burgess, 1996, cited in Balota et al., 2007). Sixteen stimuli were presented in their correct spellings (e.g., "cake"), with eight short stimuli and eight long stimuli. Half of these were HF, and half of these were LF. Sixteen stimuli were presented with incorrect spellings, with phonologically plausible errors according to the phonological representation and Spanish pronunciation (i.e., pseudo-homophones whose transcriptions followed Spanish phonological rules, e.g., "yiar" instead of "year"). To respond, participants had to press – as quickly as possible – a key on the computer keyboard (the green key if the letter sequence constituted a real word, and the red key if not). Stimuli were presented in lowercase letters (Arial, 20-point font) at the center of the screen (black on white) using DMDX software (Forster and Forster, 2003). Each stimulus remained for 4,000 ms on the screen, replaced by an asterisk as a fixation point for 500 ms, followed by a blank screen for another 500 ms. In addition, before starting, four practice trials were run to familiarize the participants with the task. The types of responses and the RTs were recorded as data. RTs were considered to be the duration between the onset of the target on the screen and the time at which the participant pressed the key. Cronbach's alpha was 0.83.

Oral and Written Semantic Classification Task

The final objective of reading is to access semantic information, which is a step toward text reading comprehension. From a logical point of view, children with dyslexia should have more difficulties in obtaining semantic information from the written word than from the orally presented words, as their principal problem is in reading. However, considering the phonological difficulties of children with dyslexia and differences between Spanish and English phonology, it would also be consistent to argue that they have inaccurate phonological-auditory representations that will make oral recognition difficult. Using this task, we wanted to ascertain whether children with dyslexia exhibited similar performance when accessing semantic information from an auditory stimulus and from a written one.

We included two modalities: oral and written presentations. For each modality, 3 semantic categories and 24 stimuli (8 per category) were selected. For the written modality, we considered the following categories: animals, body parts, and professions. For the oral modality, we considered food, clothes, and household objects. The same categories were not used in both versions (oral and written) to avoid a facilitating effect by presenting the same category twice. These kinds of categories were chosen because they receive the same levels of attention in the English textbooks for the third and fourth grades of primary education in Spain. In addition, the selected items for each category appear as part of the vocabulary in the cited English textbooks. The stimuli of the different semantic categories are matched in the number of letters (M = 4.9, SD = 1.2), phonemes (M = 3.8, SD = 0.9), syllables (M = 1.3, SD = 0.47), and lexical frequency (M = 10,947, SD = 7,715), according to the HAL frequency norms (Lund and Burgess, 1996, cited in Balota et al., 2007).

Participants received the stimulus (either written on the screen or orally by headphones), and they had to classify it as belonging to one of the three categories considered in the modality. To classify the stimuli, three pictures and one number (1, 2, and 3), associated with each category, were presented on the screen, and participants had to select the correct picture by pressing the assigned number on the keyboard (see **Figure 1**). The auditory stimuli were recorded by a 9-year-old bilingual girl with a Zoom H4N recorder and Audix Ht2-P Plantronic microphone. Subsequently, the stimuli were edited with Praat software (Boersma and Weenink, 2019). The experimental task was run on an HP Mini laptop using the DMDX program (Forster and Forster, 2003). Cronbach's alpha for the written version was 0.86, and Cronbach's alpha for the oral version was 0.72.

The research design and protocol were approved by the Ethics Committee for Research of the Principality of Asturias, Spain. The study was developed in accordance with the Declaration of Helsinki and the principles of the Spanish Law of Personal Data Protection (15/1999 and 3/2018). A written informed parental consent was received for all participants, authorizing the students to take part in the experiment (**Figure 1**).
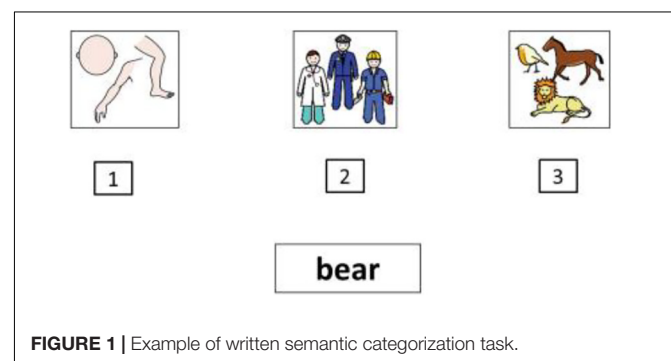

FIGURE 1 | Example of written semantic categorization task.

RESULTS

For RTs, ANOVAs were performed with mixed-effects analyses (Baayen, 2008) using R-software (R Core Team, 2016), with participants and items as random-effect variables. As fixed factors, we considered the group factor (DYS vs. CON), as well as different factors according to the task (word frequency, length, spelling type, presentation type, or type of stimuli). Stepwise model comparisons were conducted, from the most complex to the simplest model and the one with the most complex adjustment but the smallest Bayesian information criterion and the significant χ^2 test for the log-likelihood was retained (Schwarz, 1978). F values from the ANOVAs of type III, with the Satterthwaite approximation for degrees of freedom, were reported for fixed-effects variables. If interactions were significant, t tests were performed, and the p values were adjusted via the Holm–Bonferroni method. For the analysis of errors, we used a generalized mixed-effect model with a binomial distribution. A $p < 0.05$ was adopted as a level of significance.

Discrimination of Phonemes: Same–Different

In this task, we analyzed RTs and accuracy, considering group (CON vs. DYS) and stimuli type (same vs. different) as fixed-effects variables. For the analysis of RTs, we found type of stimuli effect [$F_{(1,17.578)}$ = 5.1471, $p < 0.05$), where RTs were longer for the different-stimuli pairs than for the same-stimuli pairs (estimate = 142, SE = 62.4; effect size = 0.72).

We found the same effect when accuracy was considered, with differences between same and different stimuli ($p < 0.01$; estimate = 1.72, SE = 0.55; OR = 4.95; CI = 1.47–16.63; effect-size = 0.43), as they showed a higher probability of making mistakes in different-stimuli pairs than in same-stimuli pairs. The group effect was not significant, suggesting that children with dyslexia do not have specific problems discriminating English phonemes or, alternatively, that they performed similarly to the CON group.

Reading Aloud

In the reading aloud task, we analyzed RTs and accuracy. The fixed factors were group (CON vs. DYS) and lexical frequency (HF vs. LF). We identified a group effect [$F_{(1,39.173)}$ = 9.794, $p < 0.01$], as RTs were longer in the DYS group than in the CON group (estimate = 273, SE = 91.3; effect size = 0.46). We also identified a lexical frequency effect [$F_{(1,19.911)}$ = 24.933, $p < 0.001$], as RTs were longer in LF words than in HF words (estimate = 261, SE = 55.3; effect size = 0.47).

Similar results were found when accuracy was considered (group effect: $p < 0.001$, estimate = 2.5, SE = 0.57; OR = 0.08, CI = 0.02–0.25; effect size = 0.95; and lexical frequency effect: $p < 0.001$, estimate = 2.6, SE = 0.58; OR = 0.07, CI = 0.02–0.2; effect size = 0.96). These results indicated that the DYS group showed more mistakes than the CON group. Moreover, results were better for HF words than LF words, independently of the group.

Lexical Decision-Making Task

In this task, we analyzed RTs and accuracy. The fixed factors were group (CON vs. DYS), length (short vs. long), lexical frequency (high vs. low), and spelling type (correct vs. incorrect).

We found a marginally significant group effect [$F_{(1,39.98)}$ = 3.389, $p = 0.07$, estimate = 228, SE = 124; effect size = 0.51], length effect [$F_{(1,28.63)}$ = 18.58, $p < 0.001$, estimate = 239, SE = 58.1; effect size = 0.52], and spelling type effect [$F_{(1,29.59)}$ = 6.23, $p < 0.05$, estimate = 131, SE = 52.4; effect size = 0.53). These results indicated that RTs were longer for DYS children than for CON children, for long as opposed to short stimuli, and incorrect as opposed to correct spelling stimuli.

We also found group × spelling type interaction [$F_{(1,908.22)}$ = 6.71, $p < 0.01$]. Pairwise comparison showed differences between correct and incorrect stimuli in the CON group [$p < 0.01$, $t_{(49.2)}$ = 3.744, estimate = 222.6, SE = 59.5; effect size = 0.50]. The difference between CON and DYS in the correct stimuli was marginally significant [$p = 0.09$, $t_{(45.8)}$ = 2.496, estimate = 319.5, SE = 128; effect size = 0.48). See **Figure 2**.

In accuracy, we found a group effect ($p < 0.001$, estimate = 1.29, SE = 0.24; OR = 0.21, CI = 0.105–0.425; effect size = 0.93); length effect ($p < 0.05$, estimate = 0.516, SE = 0.21; OR = 0.49, CI = 0.23–1.04; effect size = 0.71); spelling type effect ($p < 0.05$, estimate = 0.40, SE = 0.21; OR = 0.40, CI = 0.19–0.85: effect size = 0.66); and group × length × spelling type interaction ($p < 0.01$). Pairwise comparison showed differences between CON and DYS in the short correct stimuli ($p < 0.001$, estimate = 1.55, SE = 0.36; effect size = 0.87); long correct stimuli ($p < 0.01$, estimate = 1.26, SE = 0.32; effect size = 0.91); and long incorrect stimuli ($p < 0.001$, estimate = 1.68, SE = 0.32; effect size = 0.89). In addition, differences between short and long incorrect stimuli in the DYS group were marginally significant ($p = 0.08$, estimate = 0.97, SE = 0.33; effect size = 0.61). See **Figure 3**.

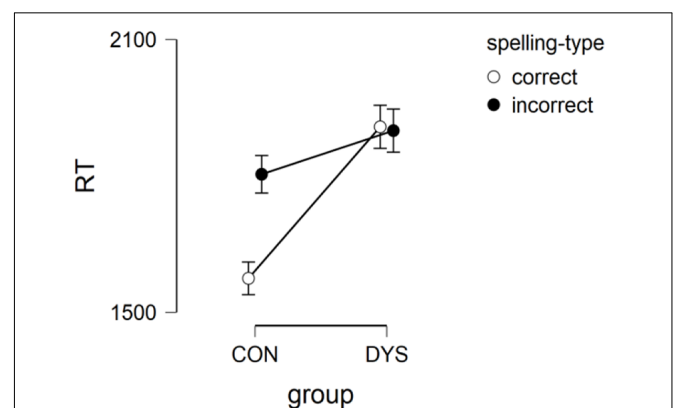Results suggested that the CON group had more orthographic representations than the DYS group.

FIGURE 2 | RT group by spelling type interaction in Lexical decision task.

Oral and Written Semantic Classification Task

We considered group (CON vs. DYS) and presentation type (oral vs. written) as fixed factors. For RTs, we found a group effect [$F(1,42.53) = 12.83$, $p < 0.001$, estimate = 830, $SE = 232$; effect size = 0.38]; presentation type [$F(1,1,528.33) = 309.541$, $p < 0.001$, estimate = 1039, $SE = 59.1$; effect size = 0.74]; and group × presentation type interaction [$F(1,1,523.13) = 55.223$, $p < 0.001$]. Pairwise comparison showed differences between groups only in written presentation ($p < 0.001$, estimate = 1,266, $SE = 238.7$; effect size = 0.36), and differences between presentation type in both DYS ($p < 0.001$, estimate = 1,475, $SE = 90.2$; effect size = 0.32) and CON ($p < 0.001$, estimate = 603, $SE = 75.7$; effect size = 0.68) groups. Children with dyslexia showed worse performance in the written version than typical readers. See **Figure 4**.

In accuracy, we found a group effect ($p < 0.001$, estimate = 1.73, $SE = 0.22$, OR = 0.11, CI = 0.06–0.19; effect size = 0.92); presentation type ($p < 0.001$, estimate = 0.51, $SE = 0.12$, OR = 0.38, CI = 0.26–0.58; effect size = 0.66); and group by presentation type interaction ($p < 0.001$). Pairwise comparison showed differences between groups in both oral presentation ($p < 0.001$, estimate = 1.29, $SE = 0.24$; effect size = 0.88) and written presentation ($p < 0.001$, estimate = 2.17, $SE = 0.26$; effect size = 0.94), but differences between presentation type only occurred in the CON group ($p < 0.001$, estimate = 0.95, $SE = 0.21$; effect size = 0.70), with more mistakes in the oral than in the written presentation.

DISCUSSION

The aim of this study was to identify specific difficulties of Spanish children with dyslexia when conducting English reading, compared to typical Spanish readers. Specifically, we tried to determine whether children with dyslexia know and use English G–P rules to read unfamiliar words or, alternatively, whether they have difficulties managing English regularities. We also tested whether they had orthographic representations of words or whether they suffered from any Spanish phonological interference. Finally, we evaluated if phonological discrimination problems were also visible in this population. To achieve our aims, four tasks were performed: discrimination of phonemes, visual lexical decision-making, reading aloud, and oral vs. written semantic categorization. Spanish children with and without dyslexia, ages 8–12, were tested.

The results suggest that Spanish children with dyslexia do not demonstrate specific problems discriminating English vowel phonemes. They performed in a similar way to children without dyslexia in terms of both RTs and accuracy. They produced better results in same-stimuli pairs compared to different-stimuli pairs. These results contradict hypotheses stating that the ability to discriminate phonemes could influence reading performance. It has been repeatedly reported that dyslexia is characterized by phonological problems, suggesting that impaired phonological or auditory processing is the origin of the reading disorder (Ahissar et al., 2000; Goswami, 2011; Peterson and Pennington, 2012). According to this view, it is possible that the stimuli or the task were not good enough to capture
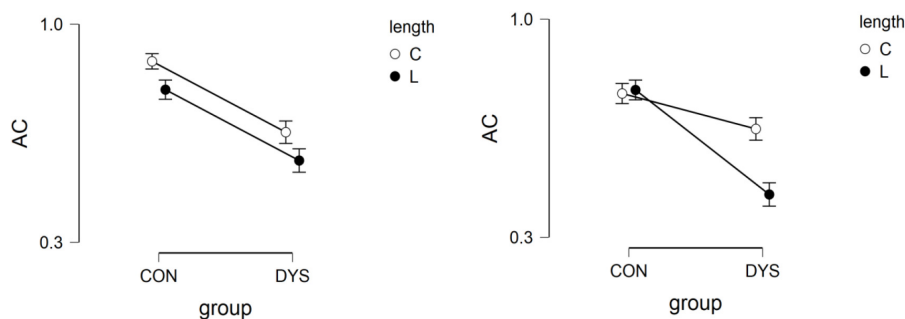

FIGURE 3 | Accuracy in group by spelling type by length interaction in Lexical decision task (**left**: correct spelling, **right**: incorrect spelling).
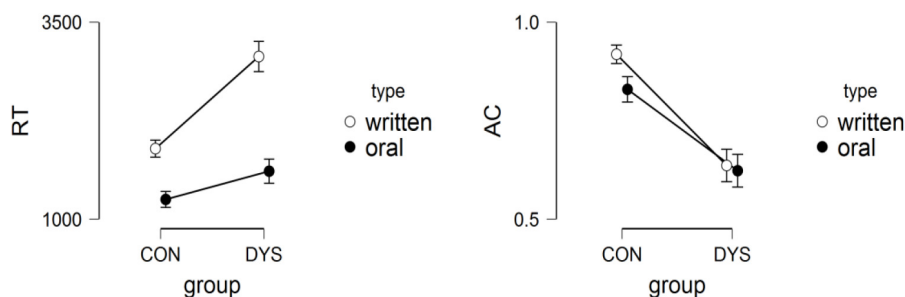

FIGURE 4 | Reaction Times (**left**) and Accuracy (**right**) in group by presentation type interaction in semantic categorization task.

the repeatedly reported phonological problems in people with dyslexia. However, alternative explanations could also be offered. First, the absence of differences could be a consequence of the age of participants, as phonological processing improves with reading experience (Perfetti et al., 1987; Morais, 1991). In this sense, another, more demanding, task would be more informative about phonological difficulties. Finally, according to the typical phonetic boundaries acquisition in L1, it is possible that children with dyslexia benefit from their difficulty in acquiring phonetic boundaries in L1, retaining the sensitivity to universal phonetic boundaries (Serniclaes et al., 2004; Soroli et al., 2010). However, it should be noted that the reliability of the phonemes discrimination task was low, so results cannot be considered as sufficiently consistent.

On the other hand, DYS children showed worse performance than CON children in all other tasks (reading aloud, visual lexical decision-making, and semantic categorization).

Considering the reading aloud task, designed to determine whether DYS children use some English G–P rules, we observed that they made more mistakes and were slower than the CON group. A similar pattern was observed in Spanish reading, although the main problem in Spanish children with dyslexia is reading slowness (Suárez-Coalla and Cuetos, 2012). However, similar to typical readers, they showed a lexical frequency effect in accuracy and RTs, so they performed better in HF than in LF words. These data seem to suggest that Spanish children with dyslexia are not able to learn G–P rules, but that they have developed orthographic representations of English words. They preferably use a lexical instead of a sublexical strategy to read (although their performance was below that of the CON children), probably given the difficulty of learning the alphabetic code. When it comes to reading in Spanish, the transparency of the orthographic system facilitates the learning of the alphabetic code, but it is not the case for the English orthographic system. In this vein, it has been reported that deep orthographic systems force people to develop lexical reading strategies (Wang et al., 2012). Our results were not in line with those of Palladino et al. (2013), who found that Italian children with dyslexia (aged 13) were accurate in reading pseudo-words. Those results were interpreted by Palladino et al. (2013) as showing that the Italian children with dyslexia have the capacity to assimilate English pronunciation rules. In our case, we used very LF words, instead of pseudo-words, so that they could potentially benefit from the pronunciation rules in reading them. However, they did not seem to benefit from these pronunciation rules, suggesting that they were using lexical reading. This idea could be consistent with the dyslexic preference for English reading hypothesis (Miller-Guron and Lundberg, 2000). Miller-Guron and Lundberg (2000) reported that some Swedish adults with dyslexia (10 in their study) prefer to read in English than in Swedish. This preference seems to start at around the fourth grade, when Swedish children begin learning English at school. At this point in time, they have already experienced a failure with the Swedish alphabetical code. The authors hypothesized that some people with dyslexia, because of their problems with learning the alphabetic code and their knowledge about English inconsistencies, either prefer or force lexical reading. This interpretation was also suggested by

Siegel et al. (1995), who argued that people with dyslexia try to compensate for the difficulty in mastering the phonemic strategy of 1:1 decoding, paying more attention to the orthographic form of English words.

To deepen our understanding of the strategies the children used during English reading, a visual lexical decision-making task was performed. In this task, real words and pseudo-homophones, whose transcriptions followed Spanish phonological rules, were included. With this task, we aimed to ascertain whether the children used a robust orthographic representation to recognize words or, alternatively, whether the Spanish phonological code affected the visual lexical decision-making task. The CON children had better performances than the DYS children, as they made fewer mistakes and were faster than the latter group. Moreover, the DYS children spent a similar amount of time on correct and incorrect stimuli, but the CON children were faster when reading real words. Finally, considering mistakes, we did not find differences between short and long stimuli in the CON group (both correct and incorrect). These data led us to conjecture that there were more robust orthographic representations in the CON group, expanding the previous data. We deduce that Spanish DYS children experience difficulties developing orthographic representations of words (Suárez-Coalla et al., 2014a,b), and they probably experience the influence of the Spanish phonological code more than typical readers.

As regards the semantic categorization task, our objective was to evaluate the possible differences between oral and written processing in DYS children. In the two previous tasks, it was not strictly necessary to know semantic information to complete the tasks. When reading aloud, the children could read words using G–P rules, and in the lexical decision-making task, they could recognize words using orthographic representations. However, in semantic categorization, they need to access the words' meanings, allowing us to compare whether they obtained semantic information from oral and written presentations in the same way. The results indicated that DYS children showed worse performance in the written version than the CON group when RTs were considered, as they spent more time than the CON group on the written stimuli. In addition, they made more mistakes than the CON group in both the oral and written versions. It should be noted, however, that typical readers benefited from the written version in terms of accuracy, while children with dyslexia did not. Considering this result, we can conclude that the DYS group also had some difficulties with English oral processing, as they made more mistakes than the CON children in the oral version. This supports the argument that the DYS group has fewer phonological representations and a smaller vocabulary than the CON group. That concurs with the suggestion that there are different problems associated with dyslexia that affect language learning (Crombie, 2000). Concretely, we suggest that weakness in phonological processing, poor working memory, and slow speed of information processing could affect performance in oral semantic categorization in particular, and language learning in general.

To summarize, a series of experiments on reading in EFL and related tasks were performed with Spanish children with dyslexia. The results suggested that Spanish children with dyslexia

demonstrate difficulties mastering English G–P rules, leading them to use a lexical strategy to read English words. However, they also demonstrated difficulties in developing orthographic representation of words, with significant consequences. Finally, the results suggested that they also show problems with oral language, demonstrating difficulties in deriving semantic information from auditory presentation.

Our results confirm previous studies on EFL reading in people with dyslexia. Previous studies have reported that English reading is a challenging task for this population. In addition, the results agree with the LCDH (Sparks et al., 1999, 2012) and subsequent studies specifically related to reading (Chodkiewicz, 1986; Durgunoglu et al., 1993; Cisero and Royer, 1995; Da Fontoura and Siegel, 1995; Geva et al., 1997; Comeau et al., 1999; Dufva and Voeten, 1999; August et al., 2001; Kahn-Horwitz et al., 2006). This supports the argument that reading problems in L1 transfer to reading in an FL, due to a common cause. In general, we confirm English reading differences between DYS and CON children, as has been previously reported (Chinese: Ho and Fong, 2005; Hebrew: Oren and Breznitz, 2005; Italian: Palladino et al., 2013; Norwegian: Helland and Kaasa, 2005; Polish: Lockiewicz and Jaskulska, 2016). However, it should be noted that our participants were younger than those of other studies (Italian: Palladino et al., 2013; Norwegian: Helland and Kaasa, 2005; Polish: Lockiewicz and Jaskulska, 2016), and the tasks were also different. In this vein, we found, contrary to the findings of Palladino et al. (2013), that Spanish children with dyslexia do not master the English G–P rules. According to Miller-Guron and Lundberg (2000), they seem to prefer a lexical strategy, but they also have problems with this strategy.

Limitations

These outcomes help us to better understand how Spanish children with dyslexia address reading in EFL. However, more evidence is needed, as reading acquisition is a very complex process. Research in this field would allow us to design strategies to improve English language teaching and learning for children with dyslexia.

Our study has limitations that should be taken into account in the future. We tried to address some of the main difficulties that Spanish children with dyslexia show when they engage in EFL reading. We wanted to identify the reading strategies that Spanish children with dyslexia use. However, the size of the group was small considering the range of ages in the sample. Therefore, the

results must be considered with caution. Furthermore, although there were no differences between the types of schools the children attended, other variables could have an important influence on our results (such as motivation, English vocabulary, and reading exposure, etc.). The findings support the argument that Spanish children with dyslexia demonstrate significant difficulties when reading in English. It is likely, however, that there are subgroups with different degrees of difficulties (perhaps affected by other variables, such as age, type of task, teaching methodology, English exposure, motivation to learn English, vocabulary level, etc.). In addition, it would be necessary to examine again the phonological awareness skills, as the task performed in this study was not sufficiently reliable. Finally, considering our results, other areas should be pursued, and a longitudinal study could contribute to greater knowledge about EFL acquisition in Spanish children with dyslexia.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee for Research of the Principality of Asturias, Spain. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

PS-C and AC contributed to the conception and design of the study. AC and CM-G collected and organized the database. PS-C performed the statistical analysis and wrote the manuscript. All authors read and approved the submitted version.

FUNDING

REFERENCES

Adams, M. J. (1990). *Beginning to Read: Thinking and Learning about Print*. Cambridge, MA: MIT Press.

Afonso, O., Suárez-Coalla, P., and Cuetos, F. (2019). Writing impairments in Spanish children with developmental dyslexia. *J. Learn. Disabil.* doi: 10.1177/0022219419876255 [Epub ahead of print].

Ahissar, M., Protopapas, A., Reid, M., and Merzenich, M. M. (2000). Auditory processing parallels reading abilities in adults. *Proc. Nat. Acad. Sci. U.S.A.* 97, 6832–6837. doi: 10.1073/pnas.97.12.6832

Álvarez-Cañizo, M., Suárez-Coalla, P., and Cuetos, F. (2018). The role of sublexical variables in development of reading fluency among Spanish children. *J. Child Lang.* 45, 858–877. doi: 10.1017/S0305000917000514

August, D. A., Calderon, M., and Carlo, M. (2001). The transfer of skills from Spanish to English. *NABE News* 1:42.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge, MA: Cambridge University Press.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behav. Res. Methods* 39, 445–459. doi: 10.3758/BF03193014

Beattie, R., and Manis, F. (2012). Rise time perception in children with reading and combined reading and language difficulties. *J. Learn. Disabil.* 46, 200–209. doi: 10.1177/0022219412449421

Boersma, P., and Weenink, D. (2019). *Praat: Doing Phonetics by Computer (Version6.0.43)*.

Boets, B., Wouters, J., van Wieringen, A., and Ghesquière, P. (2006). Auditory temporal information processing in preschool children at family risk for dyslexia: relations with phonological abilities and developing literacy skills. *Brain Lang.* 97, 64–79. doi: 10.1016/j.bandl.2005.07.026

Bonifacci, P., Canducci, E., Gravagna, G., and Palladino, P. (2017). English as a Foreign language in bilingual language-minority children, children with dyslexia and monolingual typical readers. *Dyslexia* 23, 181–206. doi: 10.1002/dys.1553

Bowey, J. A. (1995). Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. *J. Educ. Psychol.* 87, 476–487. doi: 10.1037/0022-0663.87.3.476

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Dev. Psychol.* 28, 874–886. doi: 10.1037/0012-1649.28.5.874

Burani, C., Marcolini, S., and Stella, G. (2002). How early does morpho-lexical reading develop in readers of a shallow orthography? *Brain Lang.* 81, 568–586. doi: 10.1006/brln.2001.2548

Cao, F., Bitan, T., Chou, T.-L., Burman, D. D., and Booth, J. R. (2006). Deficient orthographic and phonological representations in children with dyslexia revealed by brain activation patterns. *J. Child Psychol. Psychiatry* 47, 1041–1050. doi: 10.1111/j.1469-7610.2006.01684.x

Casalis, S., Commissaire, E., and Duncan, L. G. (2015). Sensitivity to morpheme units in English as L2 word recognition. *Writ. Syst. Res.* 7, 186–201. doi: 10.1080/17586801.2014.976165

Catts, H. W., Fey, M. E., Zhang, X., and Tomblin, J. B. (1999). Language basis of reading and reading disabilities: evidence from a longitudinal investigation. *Sci. Stud. Read.* 3, 331–361. doi: 10.1207/s1532799xssr0304_2

Chodkiewicz, H. (1986). *O Sprawnos'ci Czytania w Nauczaniu jęZyka Obcego [A Reading Skill in Foreign Language Teaching]*. Warszawa: Wydawnictwa Szkolne I Pedagogiczne.

Chung, K. K. H., and Ho, C. S. H. (2010). Second language learning difficulties in Chinese children with dyslexia: what are the reading-related cognitive skills that contribute to English and Chinese word reading? *J. Learn. Disabil.* 43, 196–211. doi: 10.1177/0022219409345018

Cisero, C. A., and Royer, J. M. (1995). The development of cross-language transfer of phonological awareness. *Contemp. Educ. Psychol.* 20, 275–303. doi: 10.1006/ceps.1995.1018

Clements-Stephens, A., Materek, A., Eason, S., Scarborough, H., Pugh, K., Rimrodt, S., et al. (2012). Neural circuitry associated with two different approaches to novel word learning. *Dev. Cogn. Neurosci.* 2(Suppl. 1), S99–S113. doi: 10.1016/j.dcn.2011.06.001

Coltheart, M., and Rastle, K. (1994). Serial processing in reading aloud: evidence for dual-route models of reading. *J. Exp. Psychol. Human.* 20, 1197–1211. doi: 10.1037/0096-1523.20.6.1197

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psycholog. Rev.* 108, 204–256. doi: 10.1037/0033-295x.108.1.204

Comeau, L., Cormier, P., Grandmaison, E., and Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *J. Educ. Psychol.* 91, 29–43. doi: 10.1037/0022-0663.91.1.29

Commissaire, E. (2012). *Orthographic and Phonological Coding During L2 Visual Word Recognition in L2 Learners: Lexical and Sublexical Mechanisms*. Villeneuve-d'Ascq: Université Charles de Gaulle - Lille III.

Constantinidou, M., and Stainthorp, R. (2009). Phonological awareness and reading speed deficits in reading disabled Greek-speaking children. *Educ. Psychol.* 29, 171–186. doi: 10.1080/01443410802613483

Crombie, M. A. (2000). Dyslexia and the learning of a foreign language in school: where are we going? *Dyslexia* 6, 112–123. doi: 10.1002/(sici)1099-0909(200004/06)6:2<112::aid-dys151>3.0.co;2-d

Cuetos, F., Martínez-García, C., and Suárez-Coalla, P. (2018). Prosodic perception problems in Spanish dyslexia. *Sci. Stud. Read.* 22, 45–51. doi: 10.1080/10888438.2017.1359273

Cuetos, F., Rodríguez, B., Ruano, E., and Arribas, D. (2007). *PROLEC-R: Batería de Evaluación de los Procesos Lectores [PROLEC-R: Battery of Assessment of Reading Processes, Revised]*. Madrid: TEA.

Cuetos, F., and Suárez-Coalla, P. (2009). From grapheme to word in reading acquisition in Spanish. *Appl. Psycholinguist.* 30, 583–681. doi: 10.1017/S0142716409990038

Cunningham, A. E. (2006). Accounting for children's orthographic learning while reading text: do children self-teach? *J. Exp. Child Psychol.* 95, 56–77. doi: 10.1016/j.jecp.2006.03.008

Da Fontoura, H. A., and Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese–English Canadian children. *Read. Writ.* 7, 139–153. doi: 10.1007/BF01026951

Davies, R., Cuetos, F., and González-Seijas, R. (2007). Reading development in a transparent orthography. *Ann. Dyslexia* 57, 179–198. doi: 10.1007/s11881-007-0010-1

Davies, R., Rodríguez-Ferreiro, J., Suárez, P., and Cuetos, F. (2013). Lexical and sub-lexical effects on accuracy, reaction time and response duration: impaired and typical word and pseudoword reading in a transparent orthography. *Read. Writ.* 26, 721–738. doi: 10.1007/s11145-012-9388-1

De Jong, P. F., and van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Sci. Stud. Read.* 6, 51–77. doi: 10.1207/S1532799XSSR0601_03

Dijkstra, A., and van Heuven, W. J. B. (2002). The architecture of the bilingual Word recognition system: from identification to decision. *Biling. Lang. Cogn.* 5, 175–197. doi: 10.1017/S136672890200301

Dufva, M., and Voeten, M. J. M. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Appl. Psycholinguist.* 20, 329–348. doi: 10.1017/S014271649900301X

Durgunoglu, A. Y., Nagy, W. E., and Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *J. Educ. Psychol.* 85, 453–465. doi: 10.1037/0022-0663.85.3.453

Ehri, L. C. (1987). Learning to read and spell words. *J. Read. Behav.* 19, 5–31. doi: 10.1080/10862968709547585

Ehri L. C. (1998). "Grapheme-phoneme knowledge is essential for learning to read words in English," in *Word Recognition in Beginning Literacy*, eds J. Metsala, and L. Ehri, (Mahwah, NJ: Erlbaum), 3–40.

Ehri, L. C., and Roberts, K. T. (1979). Do beginners learn printed words better in contexts or in isolation? *Child Dev.* 50, 675–685. doi: 10.2307/1128932

Ehri, L. C., and Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Read. Writ.* 7, 295–326. doi: 10.1007/BF03162082

Elbro, C. (1996). Early linguistic abilities and reading development: a review and a hypothesis. *Read. Writ.* 8, 453–485. doi: 10.1007/BF00577023

Elston-Güttler, K. E., Paulmann, S., and Kotz, S. A. (2005). Who's in control? Proficiency and L1 influence on L2 processing. *J. Cogn. Neurosci.* 17, 1593–1610. doi: 10.1162/089892905774597245

Forster, K. I., and Forster, J. C. (2003). DMDX: a Windows display program with millisecond accuracy. *Behav. Res. Methods Instr. Comput.* 35, 116–124. doi: 10.3758/BF03195503

Geva, E., Wade-Woolley, L., and Shany, M. (1997). Development of reading efficiency in first and second language. *Sci. Stud. Read.* 1, 119–144. doi: 10.1207/s1532799xssr0102_2

Goswami, U. (2002). Phonology, reading development, and dyslexia: a cross-linguistic perspective. *Ann. Dyslexia* 52, 141–163.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10. doi: 10.1016/j.tics.2010.10.001

Goswami, U., and Bryant, P. (1990). *Phonological Skills and Learning to Read*. Hove: Erlbaum.

Goswami, U., Ziegler, J. C., Dalton, L., and Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *J. Mem. Lang.* 45, 648–664. doi: 10.1006/jmla.2001.2790

Grainger, J., Bouttevin, S., Truc, C., Bastien, M., and Ziegler, J. (2003). Word superiority, pseudoword superiority, and learning to read: a comparison of dyslexic and normal readers. *Brain Lang.* 87, 432–440. doi: 10.1016/S0093-934X(03)00145-7

Helland, T., and Kaasa, R. (2005). Dyslexia in English as a second language. *Dyslexia* 11, 41–60. doi: 10.1002/dys.286

Ho, C. S.-H., and Fong, K.-M. (2005). Do Chinese dyslexic children have difficulties learning English as a second language? *J. Psycholinguist. Res.* 34:603618. doi: 10.1007/s10936-005-9166-1

Hogaboam, T. W., and Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. *J. Educ. Psychol.* 70, 717–729. doi: 10.1037/0022-0663.70.5.717

Jared, D., and Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *J. Mem. Lang.* 44, 2–31. doi: 10.1006/jmla.2000.2747

Kahn-Horwitz, J., Shimron, J., and Sparks, R. (2006). Weak and strong novice readers of English as a foreign language: effects of first language and socioeconomic status. *Ann. Dyslexia* 56, 161–185. doi: 10.1007/s11881-006-0007-1

Kyte, C. S., and Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *J. Exp. Child Psychol.* 93, 166–185. doi: 10.1016/j.jecp2005.09.003

Lallier, M., and Carreiras, M. (2017). Cross-linguistic transfer in bilinguals reading in two alphabetic orthographies: the grain size accommodation hypothesis. *Psychon. Bull. Rev.* 25, 386–401. doi: 10.3758/s13423-017-1273-0

Lockiewicz, M., and Jaskulska, M. (2016). Difficulties of Polish students with dyslexia in reading and spelling in English as L2. *Learn. Individ. Diff.* 51, 256–264. doi: 10.1016/j.lindif.2016.08.037

Lund, K., and Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behav. Res. Methods Instr. Comput.* 28, 203–208. doi: 10.3758/BF03204766

Lyon, G. R., Shaywitz, S. E., and Shaywitz, B. A. (2003). Defining dyslexia, comorbidity, teachers' knowledge of language and reading: a definition of dyslexia. *Ann. Dyslexia* 53, 1–14. doi: 10.1007/s11881-003-0001-9

Maloney, E., Risko, E. F., O'Malley, S., and Besner, D. (2009). Tracking the transition from sublexical to lexical processing: on the creation of orthographic and phonological lexical representations. *Q. J. Exp. Psychol.* 62, 858–867. doi: 10.1080/17470210802578385

Manis, F. R. (1985). Acquisition of word identification skills in normal and disabled readers. *J. Educ. Psychol.* 77, 78–90. doi: 10.1037/0022-0663.77.1.78

Marks, J. (2007). *English Pronunciation in Use Elementary.* Cambridge: Cambridge University Press.

Martens, V. E. G., and de Jong, P. F. (2008). Effects of repeated reading on the length effect in word and pseudoword reading. *J. Res. Read.* 31, 40–54. doi: 10.1111/j.1467-9817.2007.00360.x

Martínez-García, C., Suárez-Coalla, P., and Cuetos, F. (2019). Development of orthographic representations in Spanish children with dyslexia: the influence of previous semantic and phonological knowledge. *Ann. Dyslexia* 69, 186–203. doi: 10.1007/s11881-019-00178-6

Miller-Guron, L., and Lundberg, I. (2000). Dyslexia and second language reading: a second bite at the apple? *Read. Writ.* 12, 41–61. doi: 10.1023/A:1008009703641

Morais, J. (1991). "Constraints on the development of phonemic awareness," in *Phonological Processes in Literacy: A Tribute to Isabelle,* eds S. A. Brady, D. P. Shankweiler, and Y. Liberman, (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc), 5–27.

Nijakowska, J. (2010). *Dyslexia in the Foreign Language Classroom.* Bristol: Multilingual Matters.

Oren, R., and Breznitz, Z. (2005). Reading processes in L1 and L2 among dyslexic as compared to regular bilingual readers: behavioral and electrophysiological evidence. *J. Neuroling.* 18, 127–151. doi: 10.1016/j.jneuroling.2004.11.003

Palladino, P., Bellagamba, I., Ferrari, M., and Cornoldi, C. (2013). Italian children with dyslexia are also poor in reading English words, but accurate in reading English pseudowords. *Dyslexia* 19, 165–177. doi: 10.1002/dys.1456

Perfetti, C. A., Beck, I., Bell, L. C., and Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: a longitudinal study of first grade children. *Merrill Palmer Q. J. Dev. Psychol.* 33, 283–319.

Peterson, R. L., and Pennington, B. F. (2012). Developmental dyslexia. *Lancet* 379, 1997–2007. doi: 10.1016/S0140-6736(12)60198-6

Protopapas, A. (2007). CheckVocal: a program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behav. Res. Methods* 39, 859–862. doi: 10.3758/BF03192979

R Core Team (2016). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Ramos, J. L., and Cuetos, F. (2005). *Batería de Evaluación de los Procesos Lectores en el Alumnado del Tercer Ciclo de Educación Primaria y Educación Secundaria Obligatoria (PROLEC-SE) [Battery of Assessment of Reading Processes, for Fifth-Year Primary Education Students Onward (PROLEC-SE)].* Madrid: TEA Ediciones.

Reitsma, P. (1983). Printed word learning in beginning readers. *J. Exp. Child Psychol.* 36, 321–339. doi: 10.1016/0022-0965(83)90036-X

Reitsma, P. (1989). "Orthographic memory and learning to read," in *NATO Advanced Science Institutes series. Series D: Behavioural and social sciences, Vol. 52. Reading and Writing Disorders in Different Orthographic Systems*, eds P. G. Aaron, and R. M. Joshi, (New York, NY: Kluwer Academic/Plenum Publishers), 51–73. doi: 10.1007/978-94-009-1041-6_3

Ricketts, J., Nation, K., and Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Sci. Stud. Read.* 11, 235–257. doi: 10.1080/10888430701344306

Scanlon, D. M., Vellutino, F. R., Small, S. G., and Fanuele, D. P. (2000). Severe reading difficulties - Can they be prevented? A comparison of prevention and intervention approaches. *Paper presented at the Annual Meeting of the American Educational Research Association*, New Orleans, LA.

Schwarz, G. E. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136

Serniclaes, W., Heghe, S., Mousty, P., Carré, R., and Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *J. Exp. Child Psychol.* 87, 336–361. doi: 10.1016/j.jecp.2004.02.001

Serrano, F., and Defior, S. (2008). Dyslexia speed problems in a transparent orthography. *Ann.Dyslexia* 58, 81–95. doi: 10.1007/s11881-008-0013-6

Seymour, P. H., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cogniton* 55, 151–218. doi: 10.1016/0010-0277(94)00645-2

Share, D. L. (1999). Phonological recoding and orthographic learning: a direct test of the self-teaching hypothesis. *J. Exp. Child Psychol.* 72, 95–129. doi: 10.1006/jecp.1998.2481

Siegel, L. S., Share, D., and Geva, E. (1995). Evidence for superior orthographic skills in dyslexics. *Psychol. Sci.* 6, 250–254. doi: 10.1111/j.1467-9280.1995.tb00601.x

Snowling, M. J. (2000). *Dyslexia*, 2nd Edn. Malden: Blackwell Publishing.

Soroli, E., Szenkovits, G., and Ramus, F. (2010). Exploring dyslexics' phonological deficit III: Foreign speech perception and production. *Dyslexia* 16, 318–340. doi: 10.1002/dys.415

Sparks, R., Patton, J. M., Ganschow, L., and Humbach, N. (2012). Relationships among L1 print exposure and early L1 literacy skills, L2 aptitude, and L2 proficiency. *Read. Writ.* 25, 1599–1634. doi: 10.1007/s11145-011-9335-6

Sparks, R., Philips, L., Ganschow, L., and Javorsky, J. (1999). Students classified as LD and the college foreign language requirement: a quantitative analysis. *J. Learn. Disabil.* 32, 566–580. doi: 10.1177/002221949903200608

Stanovich, K. E., and Siegel, L. (1994). Phenotypic performance profile of children with reading disabilities: a regression-based test of the phonological core variable-difference model. *J. Educ. Psychol.* 86, 24–53. doi: 10.1037/0022-0663.86.1.24

Suárez-Coalla, P., Álvarez-Cañizo, M., and Cuetos, F. (2016). Orthographic learning in Spanish children. *J. Res. Read.* 39, 292–311. doi: 10.1111/1467-9817.12043

Suárez-Coalla, P., Avdyli, R., and Cuetos, F. (2014a). Influence of context-sensitive rules on the formation of orthographic representations in Spanish dyslexic children. *Front. Psychol.* 5:1409. doi: 10.3389/fpsyg.2014.01409

Suárez-Coalla, P., and Cuetos, F. (2012). Reading strategies in Spanish developmental dyslexics. *Ann. Dyslexia* 62, 71–81. doi: 10.1007/s11881-011-0064-y

Suárez-Coalla, P., and Cuetos, F. (2015). Reading difficulties in Spanish adults with dyslexia. *Ann. Dyslexia* 65, 33–51. doi: 10.1007/s11881-015-0101-3

Suárez-Coalla, P., Ramos, S., Álvarez-Cañizo, M., and Cuetos, F. (2014b). Orthographic learning in dyslexic Spanish children. *Ann. Dyslexia* 64, 166–181. doi: 10.1007/s11881-014-0092-5

Van Wijnendaele, I., and Brysbaert, M. (2002). Visual word recognition in bilinguals: phonological priming from the second to the first language. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 616–627. doi: 10.1037//0096-1523.28.3.616

Vellutino, F., Fletcher, J., Snowling, M., and Scanlon, D. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *J. Child Psychol. Psychiatry* 45, 2–40. doi: 10.1046/j.0021-9630.2003.00305.x

Wagner, R., and Torgesen, J. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101, 192–212. doi: 10.1037/0033-2909.101.2.192

Walley, A. C., Metsala, J. L., and Garlock, V. M. (2003). Spoken vocabulary growth: its role in the development of phoneme awareness and early reading ability. *Read. Writ.* 16, 5–20. doi: 10.1023/A:1021789804977

Wang, H.-C., Nickels, L., Nation, K., and Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Sci. Stud. Read.* 17, 369–384. doi: 10.1080/10888438.2012.749879

Wang, H.-C. H., Castles, A., and Nickels, L. (2012). Word regularity affects orthographic learning. *Q. J. Exp. Psychol.* 65, 856–864. doi: 10.1080/17470218.2012.672996

Wechsler, D. (2001). *Wechsler Intelligence Scale for Children –Revised (Spanish Adaptation)*. Madrid: TEA Ediciones.

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular reading system. *Appl. Psycholinguist.* 14, 1–33. doi: 10.1017/S0142716400010122

Wimmer, H., and Goswami, U. (1994). The influence of orthographic consistency on reading development: word recognition in English and German children. *Cognition* 51, 91–103. doi: 10.1016/0010-0277(94)90010-8

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Cross-Linguistic Word Recognition Development Among Chinese Children: A Multilevel Linear Mixed-Effects Modeling Approach

Connie Qun Guan[1,2,3]* and Scott H. Fraundorf[4]*

[1] Faculty of Foreign Studies, Beijing Language and Culture University, Beijing, China, [2] Center for the Advances of Language Sciences, University of Science and Technology, Beijing, China, [3] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, United States, [4] Department of Psychology and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, United States

The effects of psycholinguistic variables on reading development are critical to the evaluation of theories about the reading system. Although we know that the development of reading depends on both individual differences (endogenous) and item-level effects (exogenous), developmental research has focused mostly on average-level performance, ignoring individual differences. We investigated how the development of word recognition in Chinese children in both Chinese and English is affected by (a) item-level, exogenous effects (word frequency, radical consistency, and curricular grade level); (b) subject-level, endogenous individual differences (orthographic awareness and phonological awareness); and (c) their interactive effect. We tested native Chinese (Putonghua)-speaking children (n = 763) in grades 1 to 6 with both Chinese character and English word identification (lexical) decision tasks. Our findings show that (a) there were effects of both word frequency and age of acquisition in both Chinese and English, but these item-level effects generally weakened with increasing age; (b) individual differences in phonological and orthographic awareness each contributed to successful performance; and (c) in Chinese, item-level effects were weaker for more proficient readers. We contend that our findings can be explained by theoretical models that incorporate cumulative learning as the basis for development of item-level effects in the reading system.

Keywords: orthographic awareness, grapheme recognition development, multilevel linear mixed models, frequency, AoA, Chinese-English bilingual children

INTRODUCTION

By the end of the elementary school years, Chinese-speaking children can typically read up 2,500 Chinese characters and up to 2,000 words in English as a second language (L2 English) (NIES, 2012). Acquiring this system of lexical representations, which permits efficient word recognition, is an essential part of learning to read (Ehri, 2014; Perfetti and Stafura, 2014; Daniels and Share, 2018).

In this acquisition process, mapping lexical representation to spoken words creates a foundation for lexical and phonological processing and the subsequent acquisition of new words (Perfetti and Harris, 2013). Strong associations between orthography and phonology contribute to literacy in L1 (first language) Chinese (Guan et al., 2011, 2020) and in an L2 (Gunderson et al., 2011). However, we know little about the pattern of cross-linguistic word recognition development in both L1 Chinese and L2 English among Chinese children.

In the current study, we examine variation in the cognitive reading system for L1 and L2 word recognition development among Chinese children. We track the state of the system by estimating effects on reading performance both due to critical word properties, including frequency, consistency, and age of acquisition (AoA), and due to critical child-level development variables, including phonological awareness (PA) and orthographic awareness (OA). Our study is thus the first to examine both exogenous (item-level effects) and endogenous (individual differences) variation in psycholinguistic effects during the early years of literacy in both Chinese as L1 and English as L2.

Word Reading Development

Development reading research has employed simple tasks like word naming or lexical decision to uncover properties of the reading system in the early years of literacy acquisition. Evidence has accumulated that the average typically developing pupil is faster to respond to words that have pronunciations obeying the rules for the spelling–sound mappings of its constituent graphemes in English (e.g., Coltheart et al., 1993, 2001) or that are consistent with pronunciation of similar-looking words (e.g., Glushko, 1979; Andrews, 1982; Taraban and McClelland, 1987). Knowing what item attributes affect reading performance has motivated and constrained models about how cognitive reading processes function in English and in Chinese (Coltheart et al., 2001; Perfetti, 2007). Current theories can account for skilled reading of many languages, including both Chinese and English, and for the development of reading in English (e.g., Seidenberg and McClelland, 1989; Perfetti, 2007), but there is a need for theories that can explain reading development in languages other than English.

Thus, Davies et al. (2017) propose that developmental accounts of the reading system could be improved by observing how psycholinguistic effects vary with age. This is the challenge that we take up here. In particular, we examine two critical issues. First, although the general effects of the item-level variables mentioned above are well-established, it remains to be determined whether each of these variables also has an effect during word learning and whether these effects change with chronological age. Thus, we investigate whether item-level effects vary with grade level—or, in other words, the level of reading development.

Second, we examine whether these item-level effects are also modulated by individual differences in reading skill. Few studies have addressed both subject-level factors (such as readers' PA and OA) and item-level variables (including frequency and other orthographic or phonological features of words or characters) together to determine whether and to what extent these two levels of variables interact.

Item-Level Factors in Reading Development

Grapheme recognition is a hugely important skill for all children during primary school education (Shu and Anderson, 1997, 1999). Several psycholinguistic properties affect grapheme recognition, in part by affecting the ease of learning mappings between print and spoken word forms at the sublexical and lexical levels (Ho et al., 2003). Specifically, we focus on two properties of neighborhood structure, including orthography-to-phonology consistency (Taraban and McClelland, 1987) and frequency (Marslen-Wilson et al., 1994).

First, we know that oral reading in English is faster when there is a consistent mapping between orthographic representations and the corresponding phonology (Taraban and McClelland, 1987). DeFrancis (1989) has claimed that there is now little debate in English that highly consistent words are recognized quicker and more accurately. By comparison, it is generally believed that the correspondence between orthography and phonology in Chinese is more arbitrary than in English. Nevertheless, in Chinese, approximately 80% of characters afford some phonetic and semantic information (Shu et al., 2003). The phonetic radical gives a clue to the pronunciation, and the semantic radical gives a hint to character meaning. Thus, orthography-to-phonology *consistency* can be defined in Chinese as the ratio of the number of characters containing the same phonetic radical with the same pronunciation to the total number of characters containing that phonetic radical. Oral naming responses are faster and more accurate for words with high consistency (see examples under *Measures*), especially for low-frequency words, in both English (Seidenberg and Waters, 1985) and Chinese (Jared, 2002). This consistency effect has been interpreted as supporting a single mechanism for converting print into speech sounds based on statistical mappings between orthography and phonology that are learned in childhood. In particular, effects of consistency in Chinese imply that, in learning or developing the statistical mappings between orthography and phonology, orthographic similarity makes it easier to sound out individual words (Hsu et al., 2009).

Two other relevant word properties are its average AoA and frequency. We know that oral reading is faster when a word learned earlier (Cortese and Khanna, 2007) and when it is encountered more in daily usage (Marslen-Wilson et al., 1994).

Although there is consensus that each of these variables is relevant to word recognition, the developmental trajectories of their effects remain unclear. Several models of reading development (e.g., Zevin and Seidenberg, 2002; Johnston and Barry, 2006) predict that as young children's reading experience increases, many item-level effects should diminish. For instance, Zevin and Seidenberg's (2002) theoretical model proposes that as readers' total reading experience accumulates, the effects of early experience (i.e., AoA) should diminish in favor of more general properties of the orthography (i.e., the consistency of the orthography-to-phonology mapping).

Indeed, Davies et al. (2017), across a variety of methods, found that frequency and AoA effects diminish with increasing age. That is, as readers grew older, their performance was less affected by how common the words are in the language or by the time point at which they learned the words. By contrast, some studies revealed similar frequency effects in younger and older readers, in studies both of children (Burani et al., 2002) and of adults (Tainturier et al., 1989; Allen et al., 1991; Cohen-Shikora and Balota, 2016). Similarly, some studies have no significant differences in the AoA effect between younger and older adults in word naming (Morrison et al., 2002; Barry et al., 2006) or lexical decision (Barry et al., 2006). Indeed, other studies have even shown a more robust frequency effect in older compared to younger adult readers (Spieler and Balota, 2000; Morrison et al., 2002; Balota et al., 2004). This has led some researchers (e.g., Morrison et al., 2002; Ghyselinck et al., 2004; Murray and Forster, 2004) to conclude that the frequency and AoA effect do not diminish with growing overall experience.

These conflicting results may in part reflect methodological differences. Specifically, Cortese and Khanna (2007) observed that the AoA effect is larger in lexical decision than in word naming, supporting the interpretation that the lexical decision task emphasizes semantics (Chumbley and Balota, 1984). Here, we use the lexical decision task with a large sample size (over 700 participants and over 180,00 trials) that should give us ample power to detect any such developmental changes.

Interaction of Item-Level and Child-Level Factors

Our second major question was how word-level difficulty might interact with individual differences in reading skill. The *lexical quality hypothesis* (Perfetti, 1991; Perfetti and Hart, 2002) proposes that learning to read requires developing well-specified and precise phonological, orthographic, and semantic knowledge about words. Because phonology is automatically activated in character decoding (the *Universal Phonological Principle* of literacy; Perfetti and Harris, 2013), a key subject-level factor in developing these representations may be *PA*, the ability to perceive and manipulate sound units of a spoken language (Bruce, 1964; Liberman et al., 1974; Wagner and Torgesen, 1987). Evidence suggests that awareness of the phonological structure of word units plays a pivotal role in developing word representations in alphabetic orthographics, such as English (Bradley and Bryant, 1983), as well as logographic orthographies, such as Chinese, and other orthographies (Siok and Fletcher, 2001; see also Hu and Catts, 1998; Seymour et al., 2003). Indeed, PA during the preschool years plays a causal role in learning to read in the early school years (Bradley and Bryant, 1983; Treiman, 1985; Wagner and Torgesen, 1987).

Other language awareness skills are also important for developing high-quality lexical representation (Goswami and Bryant, 1990). Namely, *OA* refers to children's understanding of orthographic conventions used in the writing system adopted in a language (Treiman and Cassar, 1997). In Chinese, OA involves knowledge of orthographic features, including the sublexical form of radicals, that convey information about character meaning. Because character neighborhoods sharing the same radical are often semantically related, awareness of radical function may be a powerful device for the acquisition of literacy in Chinese. Indeed, Ho et al. (2003) demonstrated that various types of semantic radical knowledge, including about the position and the semantic category of semantic radicals, correlate significantly with character reading and sentence comprehension. The effects of OA are not limited to Chinese; OA also explains unique variance in reading English as L1 (Berninger et al., 1991, 2010).

However, we know little about the developmental trajectories of the influences of both PA and OA across years, nor how they interact with item-level factors. Further, in the cross-linguistic context, a key question is whether the kinds of connections that children make between phonology and orthography differ depending on the phonology of the language that is being learned and the orthographic units that this phonology makes salient. Here, we investigate how the effects of PA and OA in L1 Chinese and L2 English develop across years among primary school children, as well as how these subject-level factors interact with the item-level variable of frequency.

Present Study

Linear mixed-effects (LME) modeling permits a closer examination of these questions through item-level analysis of word and, ergo, character reading (Gilbert et al., 2011; Steacy et al., 2016). Here, we apply LME models to a large data set of lexical processing by children with Chinese characters and English words (365,760 total trials) to test item-level and subject-level factors that contribute to word recognition development in both Chinese and English. All participants are pupils from elementary schools sampled from an ongoing national-level reading assessment and intervention project in China (Guan et al., 2011, 2012, 2013, 2015, 2019). We examined the development of word recognition in children learning Chinese and English using a cross-sectional approach, examining speed and accuracy of lexical decision from the first through the sixth grade.

We applied LME modeling to examine accuracy and response time (RT) at the level of response to individual words, considering influences of both character-level properties (frequency, consistency, AoA) and subject-level properties (PA and OA), as well as the progressive change in these influences across grades. This allowed us to investigate (a) whether item-level effects on word recognition vary with age (e.g., the effects of frequency and AoA effects decrease, but consistency increases) and (b) whether item-level frequency interacts with subject-level effects. We further hypothesized that, due to limited language experience in L2, frequency might not play a role in L2 word recognition for lower graders (grades 1 to 3) and predict RT and accuracy for L2 English only for higher graders (grades 4 to 6).

We also address two limitations that may have contributed to inconsistency of results in previous studies. First, inconsistency in previous studies may result from limitations inherent in comparisons between group-level averages (e.g., of younger versus older children; Davies et al., 2017). Second, inconsistencies among previous observations may result also from limitations in

the range of ages or reading abilities sampled in previous studies (typical only or atypical only). If age-related changes are confined to specific phases of development or ability, then the age ranges in which reading is tested may have a critical influence on the nature of the item effects observed. Our study addressed both limitations by examining the effect of age as a continuous variable and including all readers regardless of ability.

MATERIALS AND METHODS

Participants

We recruited 763 students from three elementary schools in Zhejiang Province, China. All parents signed an informed consent form throughout the assessment and intervention periods from 2012 on. All participants spoke Mandarin at home as their L1.

Measures

Phonological Awareness in Chinese

Participants heard a novel character pronounced and were asked to write down the pinyin and tone. The maximum score (60) was earned by producing the correct pinyin onset, rime, and tone for each of 20 characters. The reliability coefficients of this set of measures ranged from 0.81 to 0.90.

Orthographic Awareness in Chinese

Following Guan et al. (2015), OA was measured by testing each of stroke awareness and radical knowledge. For stroke awareness (considered a cue for retrieval of Chinese characters; Flores d'Arcais, 1994), students tried to reproduce a character one stroke at a time in what they understood to be the appropriate order A maximum score (equal to 20) was earned by writing all 20 characters using the correct stroke order. For radical knowledge, a participant was first shown a novel character and then was asked to identify the constituent radicals that could make up that novel character. For example, for character "晴," the participants should select the appropriate constituent radicals "日" and "青" out of stimuli including the four semantic radicals (日, 口, 目, 月) and four phonetic radicals (青 , , 亲, 庆). The maximum score (20) could be earned by correctly identifying all radicals. The scores on these two tasks were summed to produce the OA score (maximum 40 points). The reliability coefficients of this set of measures ranged from 0.71 to 0.88.

Phonological Awareness in English

We measured English PA using the *sound oddity task* (Bradley and Bryant, 1983; James, 1996; Li et al., 2012) and same/different judgment task (Treiman and Zukowski, 1991). Both tasks were designed to test all of the three phonological levels: syllable, onset-rime, and phoneme.

The *sound oddity task* was adapted from James (1996) and Li et al. (2012). On each trial, children heard three words from an audio CD; the trios were constructed so that exactly two of the three words shared an initial phoneme (e.g., *bus*, *bun*, *rug*), a medial phoneme (e.g., *bun*, *gun*, *pin*), or a final phoneme (e.g., *hop*, *top*, *doll*). Participants were asked to identify the word with

the mismatching phoneme. Participants made their response by circling the word on a response sheet in which the corresponding grapheme of the tested phonemes was removed (e.g., _us, _un, _ug for *bus*, *bun*, *rug*). Practice trials were used to make sure the students understood the task. This task included 30 trios of words and took 1 min. The reliability was 0.90.

In the *same/different judgment task*, children were required to judge whether two words share a sound or not. The experimenter sounded out a pair of two spoken words that shared a sound at the beginning syllable (*hammer*, *hammock*), onset (*broom*, *brand*), or initial phonemes (*steak*, *sponge*), or at the shared final syllable (*compete*, *repeat*), shared rime (*spit*, *wit*), or shared final phonemes (*smoke*, *tack*). There were 10 word pairs for each of the six types mentioned above (60 total) and 80 word pairs that did not share a sound. It took students 3 min to complete this task. Reliability coefficients ranged from 0.86 to 0.89.

Orthographic Awareness in English

We used the Orthographic-Receptive Coding and Orthographic-Expressive Coding tasks (Berninger et al., 2010). For the receptive coding task, the children were exposed to either a real word (e.g., *word*) or a pseudoword (e.g., *wirf*) for 3 s, after which the word was removed from view. Children then had to judge whether the word (a) exactly matched a subsequently presented word (e.g., *werd* or *wirf*), (b) contained a given letter (e.g., *o* or *i*), or (c) contained a given letter group in exactly the same order (e.g., *ow* or *ir*). Stimulus items were designed so that correct answers could not be based solely on phonology but required attention to letters that had no phonological equivalent or that had alternative pronunciations. There were 30 sets of testing items in total. It took 3 min to complete this task. Reliability coefficients ranged from 0.70 to 0.78 for this measure.

For the Orthographic-Expressive Coding task, similar to a dictation task, the children were required to code the written words or pseudowords into temporary memory and reproduce all or parts of them in written format. There were 10 items of each of three types of reproductions: the whole word (e.g., *wirf*), a single letter in a designated position (e.g., the third letter in the word *last*), or multiple letters in designated positions (e.g., second and third letters in the word *last*). It took 5 min to administer this task. Reliability coefficients ranged from 0.81 to 0.85.

Frequency in Chinese and English

Three measures of Chinese word frequency were obtained, all from Chen and Shu (2001). These frequency values were highly correlated ($r = 0.84$ to $r = 0.95$), so we aggregated them by first z-scoring each measure to put them on a common scale and then averaging them. Doing so reduces the measure-specific variance associated with any particular measure of word frequency (Bollen, 1989). Similarly, for English frequency, we averaged[1] the Kuèera–Francis norms (Kucera and Francis, 1967) and the SUBTLEX$_{US}$ corpus (Brysbaert and New, 2009), which were also highly correlated ($r = 0.89$).

---

[1]We discovered after data collection that 10 of our 480 English words (2%) did not have word frequency information available in the SUBTLEX$_{US}$ corpus; these items were eliminated from analysis.

Lexical Decision in Chinese

To select materials for the lexical decision task, we randomly sampled 240 characters (40 from each grade level) from the curriculum, ensuring that the items were representative of the compound regularities and configurations of Chinese characters. The basic configurations include left–right (e.g., 妈), top–down (e.g., 骂), and outside–inside (e.g., 国). We defined characters as *high consistency* if the semantic radical appeared with the same pronunciation in more than 50% of characters (Shu and Anderson, 1999) and *low* if not, and we used the curricular grade level as a proxy for AoA. Another 240 pseudo-characters were created by adding, deleting, or shifting one stroke from the radicals within a legal character. The children received a practice trial to familiarize themselves with the task and then moved on the real testing session, in which they indicated whether each of the 480 characters was a real character or not, one a time; RT and accuracy were recorded by the computer.

Lexical Decision in English

To select materials for the lexical task in English, we randomly sampled 240 words (40 from each grade level) from the curriculum, ensuring that the testing items were representative of the letter–sound consistency, frequency of English words, and word reading level from each of six grades. Again, we took the curricular grade level as a proxy for AoA. Another 240 pseudo-characters were created by changing the onset, syllable, or rime of the real words; by swapping the letter orders within a word; or by changing a single letter or a cluster of letters within a word. The children received a practice trial to familiarize themselves with the task and then moved on to the real testing session, in which they judged whether each of the 480 words was real or not, one at a time; RT and accuracy were recorded by the computer.

Table 1 summarizes the descriptive statistics of all the variables.

Procedure

Participants completed all tasks in groups in their classrooms. The lexical decision tasks in both Chinese and English (20 min) were computerized, whereas all of the tasks assessing OA (20 min) and PA (15 min) were on paper. Across classrooms, we counterbalanced whether the computerized or paper tasks were presented first; the paper–pencil tasks were further counterbalanced in order. The tasks were later scored by two research assistants who had designed or familiarized with the tests; their inter-rater reliability was acceptable (all Pearson correlations above 0.85).

Analytic Strategy

We analyzed our data using LME models (Baayen et al., 2008; Davies et al., 2017), which can simultaneously account for both participant- and item-level differences. In mixed-effects models, the unit of analysis is the outcome of an individual trial rather than the average across multiple trials. We examined two dependent measures: (a) the accuracy of lexical decision, using a generalized mixed-effects model as the log odds (*logit*) of correctly judging a word, and (b) the RT (in ms) for correct lexical decisions, log-transformed to reduce positive skew.

Our fixed effects of interest included, at the item level, frequency, radical consistency (for Chinese only), and curricular grade level, and at the subject level, PA and OA. A further goal was to examine the interactions of pupil and character properties across age from grades 1 to 6. Thus, we allowed each of the effects named above to vary both linearly (i.e., a steady increase or decreases from grades 1 to 6) and quadratically (i.e., an effect strongest or weakest in the middle grades). Finally, because there is some evidence that, at least in English, frequency effects vary with reading skill (e.g., Perfetti and Hogaboam, 1975), we allowed the frequency effect to interact with our two measures of reading skill: PA and OA. We included only these interactions, for which we had *a priori* hypotheses; to avoid a combinatorial explosion of interaction terms given our large number of predictors, we did not include any higher-order interactions or other two-way interactions. Because all of our predictors except grade level were on arbitrary scales, we centered and *z*-scored them to facilitate comparison across variables. All variables (including grade level) were mean-centered to produce estimates of main effects averaging across the other variables, analogous to those from an ANOVA.

In all models, we included both participant, classroom, and item (word) random intercepts[2] to account for both participant differences and, critical to the motivation of the analysis, item differences. We adopted a model-based approach to outlier detection by fitting an initial model, eliminating observations with residuals more than three standard deviations from the mean, and then refitting each model (Baayen, 2008). This procedure identifies observations that are outliers after considering all fixed and random effects of interest.

All models were fit in R using package *lme4* (Bates et al., 2015). Fixed effects were tested using the Wald *z* test for logit models and the Sattherthwaite approximation to the *t* distribution for Gaussian models (package *lmerTest*; Kuznetsova et al., 2017), all with an $\alpha = 0.05$ criterion for significance.

RESULTS

Overall Grade-Level Differences

We first examine average performance from grade 1 to grade 6 in reduced models that included only student grade level. These models allow us to describe the overall pattern of grade-level differences, setting aside any individual differences (e.g., Peng et al., 2019), and to compare Chinese and English directly by including all observations with language as an additional predictor variable. **Table 2** and the top panel of **Figure 1** display these overall developmental differences with fewer than 0.1% of outlying observations removed. Overall performance did not significantly differ across languages, $p = 0.50$, and was close to 50%; because this was neither at floor nor ceiling, it allowed us ample room to detect effects of our variables of interest.

---

[2]We did also consider models with random slopes, but the model estimation process failed to converge. However, qualitative inspection of the parameter estimates from these models suggests that, had the random slopes been included, the principal conclusions would be unchanged.

TABLE 1 | Means and (in parentheses) standard deviations of all Chinese and English measures among all readers in each of six grades.

Measures		Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
Chinese							
N		137	111	121	115	123	114
PA	Total	14.1 (8.5)	14.7 (10.1)	16.0 (2.5)	16.8 (3.2)	17.1 (3.1)	19.2 (10.0)
OA	Total	18.8 (7.4)	21.0 (6.2)	24.0 (4.0)	25.1 (4.5)	27.0 (3.1)	27.9 (4.1)
Lexical decision	Accuracy	49% (20%)	50% (14%)	50% (9%)	55% (9%)	55% (8%)	56% (7%)
	RT (ms)	1,662 (544)	1,588 (360)	1,290 (288)	1,371 (211)	1,225 (161)	1,081 (228)
English							
N		137	111	121	115	123	114
PA	Total	14.0 (8.8)	14.5 (10.6)	15.7 (3.9)	16.8 (3.9)	17.4 (4.2)	19.1 (10.4)
OA	Total	18.6 (7.7)	20.8 (6.2)	23.9 (4.9)	25.3 (5.1)	26.7 (3.9)	28.1 (5.1)
Lexical decision	Accuracy	49% (20%)	50% (14%)	50% (9%)	55% (8%)	57% (7%)	58% (6%)
	RT (ms)	1,631 (501)	1,409 (338)	1,015 (238)	1,021 (75)	937 (161)	833 (197)

PA, phonological awareness; OA, orthographic awareness; RT, response time.

TABLE 2 | Fixed-effect estimates from mixed-effects logit model of lexical decision accuracy.

	Estimate	SE	Wald z	p
Intercept (baseline log odds of accuracy)	0.156	0.061	2.54	0.01
Language (English vs. Chinese)	0.022	0.086	0.26	0.80
Student grade level—linear effect	0.103	0.016	6.33	<0.001
Student grade level—quadratic effect	0.004	0.011	0.39	0.69
Language × student grade level—linear effect	0.021	0.003	6.13	<0.001
Language × student grade level—quadratic effect	0.005	0.002	2.27	0.02

Nevertheless, lexical decision accuracy increased from grade 1 to grade 6, as reflected by the significant linear effect of grade level. Further, a positive language × linear grade interaction indicated that this increase was especially steep for English. Lastly, a language × quadratic grade interaction indicates some departure from a linear growth rate for English.

Indeed, inspection of the means suggests an especially sharp, non-linear increase between grades 3 and 4. *Post hoc* tests using the Tukey correction for multiple comparisons (R package *emmeans*; Lenth, 2019) confirmed that this growth from grade 3 to grade 4 was the only significant year-to-year difference, in both Chinese ($p < 0.05$, all other $ps \geq 0.95$) and English ($p < 0.05$, all other $ps \geq 0.94$).

The bottom panel of **Figure 1** displays the grade-level differences for RTs to correct lexical decisions (180,231 trials for Chinese and 179,370 for English), and **Table 3** the results from the mixed-effects model with 0.8% of outlying RTs removed. Overall, RTs declined (i.e., became faster) from grade 1 to grade 6. Unlike for accuracy, there was also a main effect of language, with English words being responded to more quickly

than Chinese. Further, interactions with grade level indicated that this difference increased over time; RTs declined more quickly for English than for Chinese (linear term), although this change eventually leveled off (quadratic term).

Effects of Item-Level Variables
Accuracy

Next, we fit our main models including all of the item-level and subject-level variables of interest. Here, we fit models for Chinese and English separately because we had slightly different sets of predictors for the two languages (i.e., our measure of consistency was not generalizable to English). **Table 4** displays the results from the models of accuracy in Chinese and English with fewer than 0.01% of outlying observations removed from each model, and **Figure 2** plots model-predicted partial effects (via R package *remef*; Hohenstein and Kliegl, 2020) for each variable of interest.

We first turn our attention to the effect of item-level variables on lexical decision accuracy. The effect of word frequency (upper-left panels of **Figure 2**) showed different patterns of grade-level differences across languages: In Chinese, more frequent words were responded to more accurately across grade levels, but this effect diminished somewhat in higher grades as the less frequent words "caught up" in accuracy to the higher-frequency words. By contrast, in English, the overall main effect of word frequency was not significant; in early grades, lower-frequency words were actually recognized *better*, and a beneficial effect of word frequency emerged only in grade 5 and above.

Further, in Chinese, the word frequency effect in accuracy was qualified by interactions with both orthographic and PA such that word frequency was less important for higher-skilled readers; there were no such interactions in English. Note, however, that the standardized parameter estimates for the interactions were of substantially smaller magnitude than the main effect of frequency; that is, the frequency effect was reduced for readers of higher skill but not eliminated.

The effect of consistency in Chinese words (upper-middle panel of **Figure 2**) varied linearly across grade levels. At lower grades, low-consistency words were responded to slightly more accurately than high-consistency words, but this reversed over
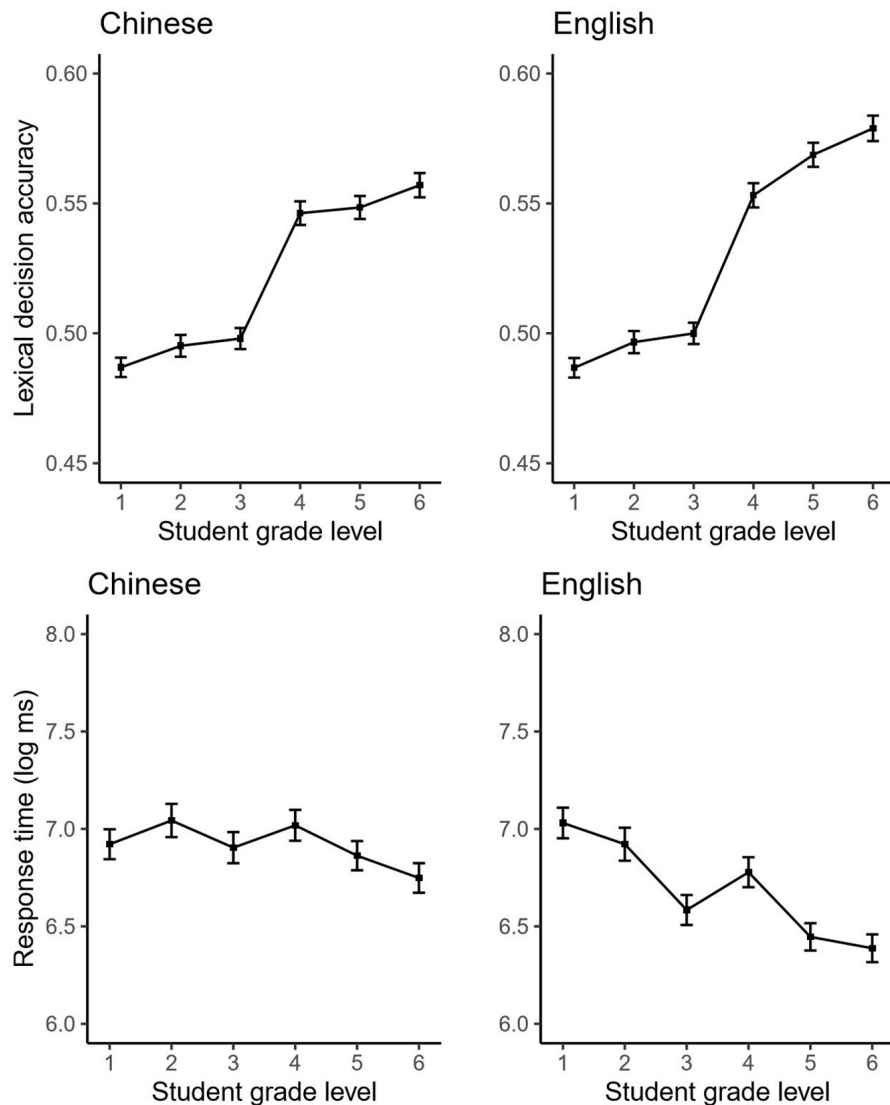
FIGURE 1 | Proportion accuracy for lexical decisions (top panels) and response time for correct lexical decisions (bottom panels) as a function of student grade level in both Chinese (left panels) and English (right panels). Error bars indicate 95% confidence intervals computed across subjects.

time such that high-consistency words were eventually judged more accurately.

Lastly, words with earlier AoA were generally responded to more accurately (upper-right panels of **Figure 2**). AoA did not have a significant main effect on accuracy in Chinese but interacted with student grade level such that the benefit of AoA was evident most strongly in earlier grades. By comparison, in English, the benefit of word AoA was strongest in middle grades, and the main effect of AoA was also significant across grades.

Response Time

Next, we turn to how these same variables affected RTs in correct lexical decision trials. **Table 5** displays the results of these models, with 0.8% and 1.1% of outlying RTs removed in Chinese and English, respectively.

Word frequency (lower-left panels of **Figure 2**) did not have a significant main effect on RTs in Chinese; there was, however, a significant developmental trend such that a frequency effect began to emerge in higher grades. By contrast, frequency had a facilitatory effect on RTs across grade levels in English, and this frequency difference increased with grade level as recognition of high-frequency words especially accelerated.

The frequency effect in Chinese was qualified by an interaction with OA such that frequency speeded responding more for students with poor OA; again, however, this interaction was of relatively small magnitude such that OA modulated but did not eliminate the frequency effect. The English frequency was also qualified by an interaction but in the opposite direction: Students with higher OA in English showed a *larger* frequency effect.

TABLE 3 | Fixed-effect estimates from mixed-effects logit model of response time for accurate lexical decisions.

	Estimate	SE	t	df	p
Intercept (baseline log RT)	6.846	0.048	142	16.58	<0.001
Language (English vs. Chinese)	−0.263	0.017	−15.19	979	<0.001
Student grade level—linear effect	−0.085	0.017	−4.92	15.78	<0.001
Student grade level—quadratic effect	−0.007	0.012	−0.63	15.71	0.54
Language × student grade level—linear effect	−0.086	0.001	−77.6	>355,000	<0.001
Language × student grade level—quadratic effect	0.026	0.001	34.8	>355,000	<0.001

Radical consistency (lower-middle panel of **Figure 2**) had no effects on RTs. Curricular grade level (lower-right panels of **Figure 2**) had significant main effects in both Chinese and English such that words with earlier AoA were responded to more quickly across grade levels. For Chinese, a significant quadratic trend indicated that this effect was largest in the middle grades, whereas for English, the effect became larger beyond the first grade.

Summary

Word frequency facilitated both the accuracy and speed of lexical decision but showed different patterns of grade-level differences across languages. The benefit of frequency on accuracy diminished with grade level in Chinese but *increased* over time in English. Nevertheless, in both languages, the benefit on RTs was largest in later grades.

The benefit of frequency was especially large for students with poor PA or OA in Chinese, whereas in English, frequency was more beneficial for students with *higher* OA.

Even when controlling for word frequency, words learned earlier in the curriculum (i.e., earlier AoA) were generally responded to more quickly and accurately. Similar to frequency, this effect was stronger in earlier grades in Chinese but stronger in later grades for English. Lastly, the consistency of Chinese radicals did not affect RT, but it did have varying effects on response accuracy, such that high-consistency words were initially responded to less accurately but, in later grades, more accurately.

Effects of Student-Level Variables

Accuracy

To analyze the student-level variables, we first return to **Table 4** to consider their effect on accuracy. PA (upper-left panels of **Figure 3**) had a main effect on accuracy in both languages such that students with greater PA responded substantially more accurately; in both languages, this effect was largest in the early grades.

The effect of OA on accuracy (upper-right panels of **Figure 3**) was even more similar across languages. Students with greater OA responded more accurately, but there were significant linear

and quadratic developmental trends in both languages, such that the effect of OA was largest in the earlier grades, smallest in the middle grades, and moderately sized in the upper grades.

Recall, further, that the benefits of OA and PA in Chinese were qualified by an interaction with word frequency such that OA and PA were most beneficial for lower-frequency words. Nevertheless, the standardized estimate for this interaction was small relative to the main effects of PA and OA; thus, PA and OA were helpful even for judging high-frequency words.

Response Time

In contrast to accuracy, PA did not have a reliable main effect on RT in Chinese (lower-left panels of **Figure 3**). However, there was a significant linear trend; PA benefited RT in earlier grades, but this effect disappeared over time. In L2 English, there was a significant main effect, but this effect nevertheless declined over time as well.

For OA (lower-right panels of **Figure 3**), there was a significant facilitatory main effect across grade levels in Chinese but no significant effects on RT in L2 English.

Summary

PA and OA had more robust effects on accuracy than RT. The developmental trend of these effects was similar across languages such that these abilities most benefited performance in the earlier grades and showed diminished effects in the higher grades. OA benefited both accuracy and RT in Chinese (with the benefit to accuracy again being largest in the earliest grades) but benefited only accuracy in L2 English.

The benefits of orthographic and PA in Chinese were stronger for lower-frequency words; that is, good PA and good OA could help compensate for the difficulty associated with reading low-frequency words.

DISCUSSION

In this current study, we explored the general development of word recognition development across grades in L1 Chinese and L2 English, as well as how these grade-level differences are influenced by both item- and subject-level characteristics. Using the lexical decision task, we assessed word recognition of 240 Chinese characters and 240 English words cross-sectionally from grade 1 to grade 6. We used LME modeling to simultaneously consider item-level (frequency, consistency, and curricular grade level) and subject-level (OA and PA) variables.

Three major findings were obtained. First, as grade level increases, accuracy increases and RT speeds up for both English and Chinese. In particular, it seems that the transition from grade 3 to grade 4 (with students' age between 10 and 11 years old) is a period when accuracy in word recognition sharply increases Second, word frequency and curricular grade level each predict word recognition in both languages but develop differently across grades, with the benefits of word frequency stronger in early grades in L1 Chinese but in later grades (i.e., grade 4 and above) in L2 English. The benefit of consistency of Chinese characters also increased with students' age from grade 1 to grade 6. Third,

TABLE 4 | Fixed-effect estimates from mixed-effects logit model of lexical decision accuracy for Chinese (top panel) and English (bottom panel) as a function of item- and student-level variables.

	Estimate	SE	Wald z	p
Chinese				
Intercept (baseline log odds of accuracy)	0.338	0.180	1.88	0.06
Student grade level—linear effect	−0.205	0.018	−11.52	<0.001
Student grade level—quadratic effect	0.015	0.012	1.28	0.20
Item-level variables				
Frequency	0.222	0.085	2.62	0.01
Frequency × linear student grade level	−0.010	0.004	−2.72	0.01
Frequency × quadratic student grade level	−0.012	0.002	−5.32	<0.001
Consistency	0.017	0.064	0.27	0.79
Consistency × linear student grade level	0.021	0.002	8.47	<0.001
Consistency × quadratic student grade level	<0.001	0.002	−0.06	0.95
Curricular grade level	−0.070	0.047	−1.50	0.13
Curricular grade level × linear student grade level	0.017	0.002	9.35	<0.001
Curricular grade level × quadratic student grade	0.005	0.001	3.99	<0.001
Student-level variables				
Phonological awareness	0.165	0.043	3.81	<0.001
Phonological awareness × linear student grade level	−0.014	0.010	−1.41	0.16
Phonological awareness × quadratic grade level	−0.016	0.009	−1.85	0.06
Orthographic awareness	0.599	0.036	16.63	<0.001
Orthographic awareness × linear student grade level	−0.034	0.013	−2.58	0.01
Orthographic awareness × quadratic grade level	0.022	0.009	2.61	0.01
Frequency × phonological awareness	−0.052	0.006	−8.19	<0.001
Frequency × orthographic awareness	−0.020	0.005	−3.77	<0.001
English				
Intercept (baseline log odds of accuracy)	0.645	0.137	4.71	<0.001
Student grade level—linear effect	−0.116	0.017	−6.66	<0.001
Student grade level—quadratic effect	0.030	0.011	2.59	0.01
Item-level variables				
Frequency	−0.035	0.060	−0.58	0.56
Frequency × linear student grade level	0.050	0.003	16.79	<0.001
Frequency × quadratic student grade level	0.014	0.002	8.13	<0.001
Curricular grade level	−0.144	0.034	−4.30	<0.001
Curricular grade level × linear student grade level	0.015	0.001	10.44	<0.001
Curricular grade level × quadratic student grade	0.005	0.001	4.84	<0.001
Student-level variables				
Phonological awareness	0.230	0.040	5.77	<0.001
Phonological awareness × linear student grade level	−0.012	0.010	−1.20	0.24
Phonological awareness × quadratic grade level	−0.027	0.009	−3.14	<0.01
Orthographic awareness	0.396	0.034	11.53	<0.001
Orthographic awareness × linear student grade level	−0.089	0.012	−7.27	<0.001
Orthographic awareness × quadratic grade level	0.026	0.008	3.13	<0.01
Frequency × phonological awareness	−0.005	0.005	−1.01	0.31
Frequency × orthographic awareness	> −0.001	0.006	−0.07	0.95

we observed item-by-subject interactions in Chinese such that both PA and OA were more beneficial to low-frequency words in accuracy; OA was also more beneficial to low-frequency words in RT. We did not observe this interaction in L2 English; if anything, OA was more beneficial for high-frequency words in L2 English.

We discuss these major results first in terms of our statistical approach. We then turn to the item-level and subject-level effects and their interaction effects and what these effects indicate about the development of word recognition. Finally, we provide some

consideration of how theoretical models of reading development generalize to a cross-linguistic perspective on word recognition.

Mixed Linear Modeling of Cross-Linguistic Developmental Data

The development of multilevel LME models permits a closer look at word recognition development through item-level analysis of word reading (e.g., Gilbert et al., 2011;
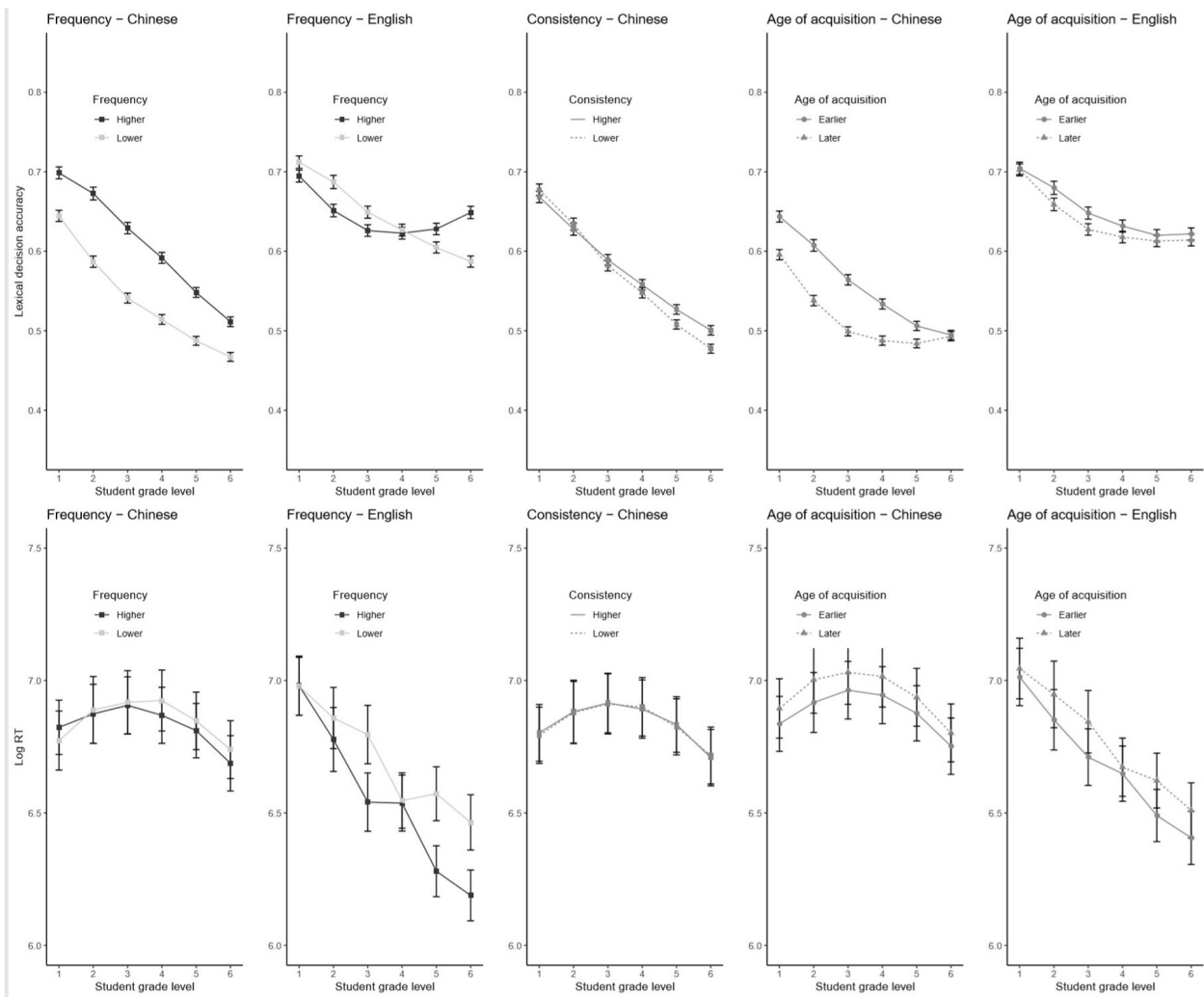
FIGURE 2 | Model-predicted proportion accuracy for lexical decisions (top panels) and response time for correct lexical decisions (bottom panels) as a function of the partial effects of student grade level and item-level properties. Frequency, consistency, and age of acquisition are depicted as median splits for purposes of visualization but were entered as continuous variables into the mixed-effects models. Error bars depict 95% confidence intervals across subjects.

Steacy et al., 2016; Guan et al., 2020). Here, we applied such models to understanding the development of word recognition from a cross-linguistic perspective. Similar to the growth curve analyses conducted in previous research (Berninger et al., 2010, 2013; Goswami, 2010), we examined how word recognition changed between grades 1 and 6—were they steady linear changes, or did they show asymptotic or other non-linear changes?

At the broadest level, the models showed similar and generalizable patterns of word learning development across languages, i.e., as grade level increases, the recognition accuracy increases and RT speeds up for both English and Chinese. In particular, for both L1 Chinese and L2 English, the recognition accuracy increased sharply from grade 3 to grade 4 but plateaued afterward.

A particular contribution of this current study is the use of mixed effects to simultaneously examine not only item- and

subject-level effects but also their interactions (and for both L1 Chinese and L2 English). We discuss those effects more in detail below.

Item-Level Effects

We found that two item-level variables—word frequency and AoA (operationalized here as curricular grade level)—were beneficial in both languages. Further, AoA showed similar grade-level differences across languages such that it diminished with advancing grade levels. Nevertheless, frequency showed somewhat different patterns across languages: In L1 Chinese, the benefit of high frequency diminished with grade level, but in L2 English, high-frequency words were initially judged less accurately, and frequency only became beneficial later.

It is noteworthy that, in general, these item-level effects decreased with age. Murray and Forster (2004) had argued that

TABLE 5 | Fixed-effect estimates from mixed-effects logit model of response time for accurate lexical decisions for Chinese (top panel) and English (bottom panel) as a function of item- and student-level variables.

	Estimate	SE	t	df	p
Chinese					
Intercept (baseline log RT)	6.912	0.064	108	28.44	<0.001
Student grade level—linear effect	−0.014	0.021	−0.69	17.79	0.50
Student grade level—quadratic effect	−0.025	0.014	−1.78	16.39	0.09
Item-level variables					
Frequency	−0.024	0.016	−1.52	493	0.13
Frequency × linear student grade level	−0.011	0.001	−9.62	>175,000	<0.001
Frequency × quadratic student grade level	0.004	0.001	5.75	>175,000	<0.001
Consistency	<0.001	0.012	0.01	490	0.99
Consistency × linear student grade level	−0.001	0.001	−0.96	>175,000	0.34
Consistency × quadratic student grade level	0.001	0.001	0.50	>175,000	0.62
Curricular grade level	0.024	0.009	2.70	491	0.01
Curricular grade level × linear student grade level	> −0.001	0.001	−0.20	>175,000	0.84
Curricular grade level × quadratic student grade	−0.001	0.001	−2.15	>175,000	0.03
Student-level variables					
Phonological awareness	−0.023	0.030	−0.76	703	0.45
Phonological awareness × linear student grade level	0.023	0.007	3.48	708	0.001
Phonological awareness × quadratic grade level	−0.002	0.006	−0.27	705	0.79
Orthographic awareness	−0.058	0.025	−2.33	709	0.02
Orthographic awareness × linear student grade level	> −0.001	0.009	−0.05	703	0.96
Orthographic awareness × quadratic grade level	−0.001	0.006	−0.21	712	0.84
Frequency × phonological awareness	<0.001	0.002	−0.27	>175,000	0.79
Frequency × orthographic awareness	0.004	0.002	2.26	>175,000	0.02
English					
Intercept (baseline log RT)	6.614	0.053	124	26.91	<0.001
Student grade level—linear effect	−0.130	0.017	−7.55	17.10	<0.001
Student grade level—quadratic effect	0.008	0.012	0.66	16.56	0.52
Item-level variables					
Frequency	−0.102	0.012	−8.61	488	<0.001
Frequency × linear student grade level	−0.032	0.001	−34.19	>175,000	<0.001
Frequency × quadratic student grade level	<0.001	0.001	0.81	>175,000	0.42
Curricular grade level	0.032	0.007	4.76	488	<0.001
Curricular grade level × linear student grade level	0.004	<0.001	9.32	>175,000	<0.001
Curricular grade level × quadratic student grade	−0.001	<0.001	−3.31	>175,000	<0.01
Student-level variables					
Phonological awareness	−0.048	0.016	−3.02	695	<0.01
Phonological awareness × linear student grade level	0.014	0.004	3.53	699	<0.001
Phonological awareness × quadratic grade level	0.005	0.003	1.42	698	0.16
Orthographic awareness	−0.021	0.014	−1.50	700	0.14
Orthographic awareness × linear student grade level	0.006	0.005	1.16	695	0.25
Orthographic awareness × quadratic grade level	−0.045	0.003	−1.29	703	0.20
Frequency × phonological awareness	<0.001	0.002	0.09	>175,000	0.93
Frequency × orthographic awareness	−0.007	0.002	−3.56	>175,000	<0.001

the frequency effect in lexical access or word recognition should not change along with growing overall experience. However, later, based on findings from a range of methods, Davies et al. (2017) suggested that word frequency and AoA effects decline with increasing age. That is, as readers grow older and gain more experience, their performance is less affected by how common the words are in the language or by the time point at which they learnt the words. This is likely because readers in more advanced grades

have encountered more of these words and thus can handle them all more accurately. Our results support this latter claim.

Within L1 Chinese, we also examined a third item-level variable: radical consistency. For this variable, we found that high consistency was associated with superior recognition in later grades but poorer performance in earlier grades. Previous literature has not provided a clear picture on the development of this consistency effect, because grade levels have been sampled
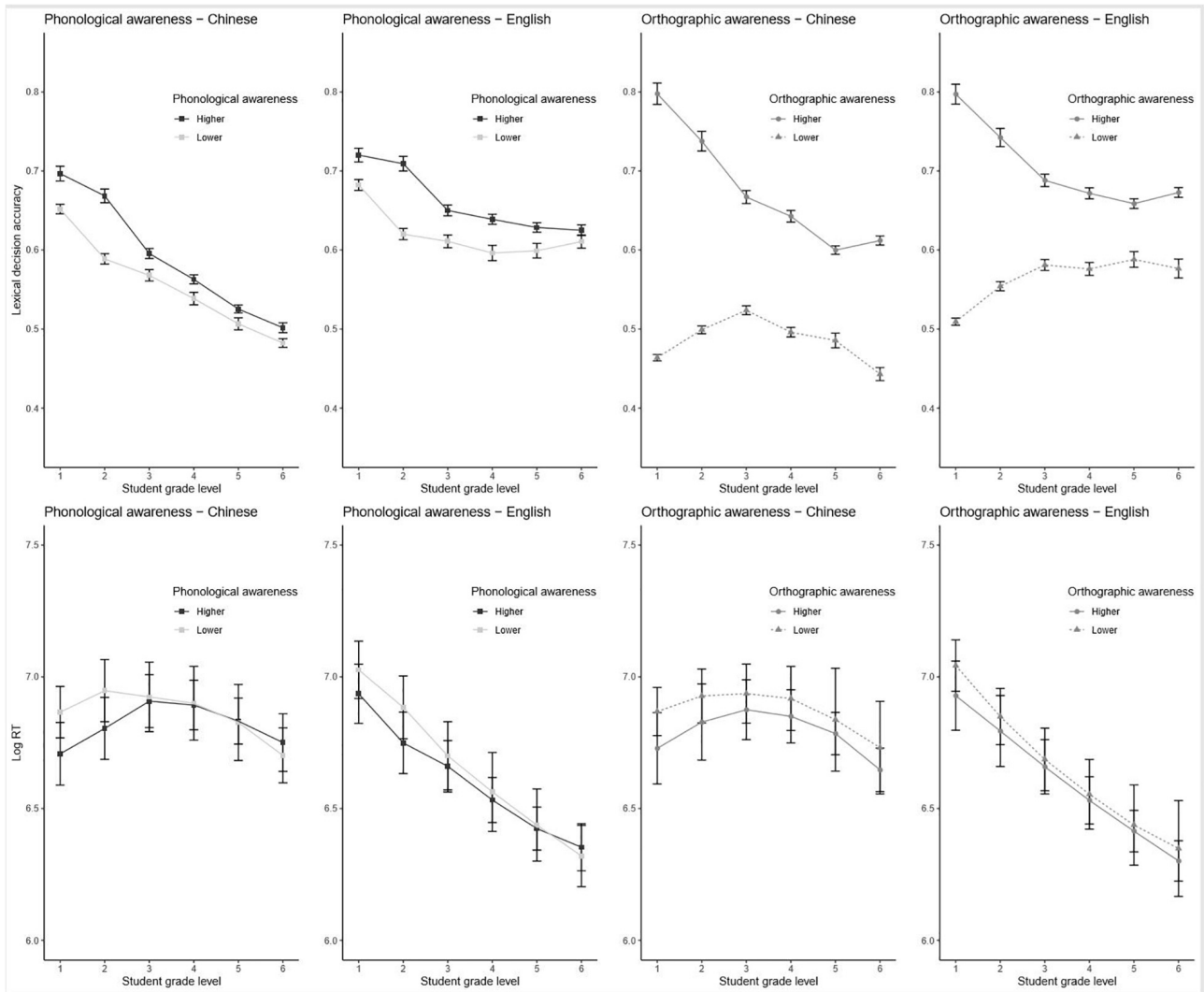
FIGURE 3 | Model-predicted proportion accuracy for lexical decisions (top panels) and response time for correct lexical decisions (bottom panels) as a function of the partial effects of student grade level and subject-level properties. Phonological and orthographic awareness are depicted as median splits within each grade level for purposes of visualization but were entered as continuous variables into the mixed-effects models. Error bars depict 95% confidence intervals across subjects.

somewhat sporadically. For example, Yang and Peng (1997) tested third- and sixth-grade school children in a naming task and found that both showed a consistency effect (as defined in Fang et al., 1986). Shu and Wu (2006) replicated the experiment of Yang and Peng (1997) with fourth- and sixth-grade children and found that both showed consistency effects. Shu et al. (2000) found that this effect grew stronger as children got older. Shu et al. (2003) have also found that children need a long time to develop phonetic consistency awareness. Our results are also consistent with this claim in that we found that consistency was only beneficial in later grades.

Taken together, our results suggest continuous development of word learning in both Chinese and English. The developmental patterns begin at an earlier age in L1 Chinese and at a later age in L2 English. A plausible interpretation is that the effects of word features like frequency and consistency begin to manifest after the learners have grasped some basic awareness and knowledge of word-level skills—at middle grades (e.g., grade 3) in L1 Chinese and advanced grades in L2 English (e.g., grades 5 and 6), since English is introduced in formal classroom instruction after grade 3 (NIES, 2012). Interestingly, these item-level effects may interact with subject-level effects, which we discuss below.

Subject-Level Effects

The subject-level effects suggest a general benefit of PA and OA in word recognition, though mainly in response accuracy rather than RT. The benefits of these skills were largest in earlier grades, when beginning readers may not yet have other applicable skills or knowledge. These findings are consistent with prior work, so the subject-level effects alone are not a major focus in the current study.

Interaction Effects

Of greater interest was how the subject-level factors moderated the strength of item-level effects. PA and OA interacted with character frequency in L1 Chinese to affect response accuracy, and in the case of OA, it interacted with character frequency to affect RT. Specifically, readers with lower PA and OA benefited more from character frequency, whereas readers with high skill could handle even low-frequency characters in Chinese. To put it another way, reading skill mattered more when reading low-frequency characters than high-frequency ones. This is consistent with past evidence that frequency effects are generally larger for less-skilled readers (e.g., Perfetti and Hogaboam, 1975; Davies et al., 2017); here, we show that these effects extend to developing L1 Chinese readers.

In contrast, there were no frequency × PA interactions in L2 English, and the frequency × OA interaction was reversed such that students with higher OA in English showed a larger frequency effect. We suspect that this might be due to the fact that language experience differs between Chinese and English in our sample. In this study, we recruited students who were beginning learners of English as an L2, i.e., they were not balanced Chinese–English bilinguals. These students were just beginning to accumulate their language experience, such that only those students with higher OA may have been able to capitalize on word frequency. That is, even those students relatively high in L2 English OA may have only had a level of reading ability equal to what constituted "poor" OA in L1 Chinese.

A Theoretical Model of Reading Development Generalizable Across Languages

The theoretical model of Zevin and Seidenberg (2002) predicts that effects of consistency, frequency, and word AoA vary over time. As readers accumulate experience, their initial experiences (i.e., AoA) matter less, and their performance becomes instead dominated by more general regularities of the orthography-to-phonology mapping. Although our goal was not to conduct a global and complete test of this model, we at least provide supportive evidence by showing that (a) AoA effects diminish across grades, whereas (b) effects of a radical's phonetic consistency become larger.

The interactions of age with frequency or AoA are consistent with a gradual ceiling effect predicted to result from the assumption—inherent in connectionist network systems—of asymptotic learning based on distributed representations and a non-linear input–output function (Van Orden et al., 1990; Plaut et al., 1996). That is to say, the effects of psycholinguistic properties change as a function of the oral reading system, approaching maximal efficiency as experience accumulates and skill develops. Another example of this principle is that, while the consistency effect in English influences children's reading (Laxon et al., 1988, 2002), it is smaller for more skilled readers (Laxon et al., 1988).

This is because the other reading component skills, such as PA or OA, develop and compensate for difficult words. We observed similar effects in our study insofar as frequency effects were weaker in L1 Chinese for readers high in PA and/or OA.

This principle of asymptotic word learning applies cross-linguistically in both L1 Chinese and L2 English. For instance, in the present study, we found that AoA effects diminished with grade level increases in both L1 Chinese and L2 English. Indeed, these features of connectionist reading models can apply to all languages and any type of script provided that the statistical constraints of a specific language are known beforehand.

Future Directions

In this study, we conducted a cross-sectional comparison of grades 1 and 6. At an empirical level, future studies could examine the developmental patterns of cross-linguistic word learning across even broader sections of the life span and could collect longitudinal, rather than cross-sectional, data. Davies et al. (2017) argue that frequency effects change with age, most principally in the transition from childhood into adulthood. In their item-level analysis, the frequency effect was larger in children's RTs than in young adults.' In their subject-level analyses, the per-subject estimates of the frequency effect coefficient varied in relation to age, but the age effect on frequency coefficients was curvilinear; it appeared to be stronger for younger children.

At a technical level, we encourage future researchers to consider the use of an LME model to assess word learning and reading development across the life span. Researchers have typically focused either on the effects of word properties in item-level analyses or on the effects of individual differences in subject-level analyses. The benefit of a multilevel analysis of reading, such as ours, is that it allowed for the examination of item-by-subject interactions. One insight from this approach is that the psycholinguistic effects of Chinese characters on the development of literacy systematically vary in relation to individual differences in age and reading ability of a pupil. Second, variation in stimulus properties emerges against a backdrop of large, overarching, effects on performance due to individual differences. Mixed-effects models show that the effects of word properties, and their modulation by individual differences, are significant, but that the dominant source of variance in reading performance is those individual differences (see Davies et al., 2017).

Lastly, more comparable language-specific measures for both Chinese and English should be designed and validated. We analyzed Chinese and English in separate models because we did not have a comparable measure of one item-level variable, consistency, for English, which would have allowed us to directly compare languages within a single model. Determining English consistency would require hand calculation (e.g., Weekes et al., 2006); this was outside the scope of the current study but could be conducted in the future for more comparable models. There were also some limitations in the measures we did obtain. For instance, our expressive coding task in English also required

children to hold material in working memory, so variation in these scores might reflect memory skills as well as orthographic skills. Similarly, one of our Chinese OA tasks, radical knowledge, could potentially be solved on the basis of visual analysis alone—but note that this was not true of the other task measuring OA in Chinese, stroke awareness.

CONCLUSION

Our study shows the importance of both stimulus-related, item-level (exogenous), and individual-related, child-level (endogenous), psycholinguistic factors in learning to recognize words. First, we found similar trends for word reading development in both L1 Chinese and L2 English in a cross-sectional comparison of Chinese elementary students from grades 1 to 6, and we assume that this serves as a proxy for age-related effects. Second, and most importantly, we contribute evidence that the constraints on acquisition of literacy in Chinese as an L1 and English as an L2 are multifaceted and include exogenous (stimulus-related) properties as well as endogenous (subject-related) properties. We conclude that these properties interact to produce literacy in Chinese and English and form the generalizable basis of a theoretical view of early-years reading from the cross-linguistic perspective.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Science and Technology Beijing. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

CG conceived, designed, and performed the experiments. CG and SF analyzed the data and wrote the manuscript.

FUNDING

REFERENCES

Allen, P. A., Madden, D. J., and Crozier, L. C. (1991). Adult age differences in letter-level and word-level processing. *Psychol. Aging* 6, 261–271.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Mem. Cogn.* 10, 565–575.

Baayen, R. H. (2008). *Analyzing Linguistic data: A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., and Yap, M. (2004). Visual word recognition of single-syllable words. *J. Exp. Psychol. Gen.* 133, 283–316.

Barry, C.,. Johnston, R. A., and Wood, R. F. (2006). Effects of age of acquisition, age, and repetition priming on object naming. *Vis. Cogn.* 13, 911–927.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48.

Berninger, V. W., Abbott, R. D., Nagy, W., and Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *J. Psycholinguist. Res.* 39, 141–163. doi: 10.1007/s10936-009-9130-6

Berninger, V. W., Lee, Y. L., Abbott, R. D., and Breznitz, Z. (2013). Teaching children with dyslexia to spell in a reading-writers' workshop. *Ann. Dyslexia* 63, 1–24. doi: 10.1007/s11881-011-0054-0

Berninger, V. W., Mizokawa, D. T., and Bragg, R. (1991). Scientific practitioner: theory-based diagnosis and remediation of writing disabilities. *J. Sch. Psychol.* 29, 57–79.

Bollen, K. A. (1989). *Structural Equations With Latent Variables*. New York, NY: John Wiley & Sons.

Bradley, L., and Bryant, P. E. (1983). Categorizing sounds and learning to read: a causal connection. *Nature* 301, 419–421.

Bruce, D. J. (1964). The analysis of word sounds by young children. *Br. J. Educ. Psychol.* 34, 158–170.

Brysbaert, M., and New, B. (2009). Moving beyond Kuèera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav. Res. Methods* 41, 977–990. doi: 10.3758/BRM.41.4.977

Burani, C., Marcolini, S., and Stella, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain Lang.* 81, 568–586.

Chen, H. C., and Shu, H. (2001). Lexical activation during the recognition of Chinese characters: evidence against early phonological activation. *Psychon. Bull. Rev.* 8, 511–518.

Chumbley, J. I., and Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Mem. Cogn.* 12, 590–606.

Cohen-Shikora, E. R., and Balota, D. A. (2016). Visual word recognition across the adult lifespan. *Psychol. Aging* 31, 488–502. doi: 10.1037/pag0000100

Coltheart, M., Curtis, B., Atkins, P., and Haller, M. (1993). Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychol. Rev.* 100, 589–608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256.

Cortese, M. J., and Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Q. J. Exp. Psychol.* 60, 1072–1082.

Daniels, P. T., and Share, D. L. (2018). Writing system variation and its consequences for reading and dyslexia. *Sci. Stud. Read.* 22, 101–116.

Davies, R. A., Arnell, R., Birchenough, J. M., Grimmond, D., and Houlson, S. (2017). Reading through the life span: individual differences in psycholinguistic effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 43, 1298–1338. doi: 10.1037/xlm0000366

DeFrancis, J. (1989). *Visible Speech: The Diverse Oneness of Writing Systems*. Honolulu, HI: University of Hawaii Press.

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory and vocabulary learning. *Sci. Stud. Read.* 18, 5–21.

Fang, S. P., Horng, R. Y., and Tzeng, O. J. L. (1986) "Consistency effects in the Chinese character and pseudo-character naming tasks," in *Linguistics, Psychology, and the Chinese Language*, eds H. S. R. Kao and R. Hoosain (Hong Kong: Centre of Asian Studies), 11–21.

Flores d'Arcais, G. B. (1994). Order of strokes writing as a cue for retrieval in reading Chinese characters. *Eur. J. Cogn. Psychol.* 6, 337–355.

Ghyselinck, M., Lewis, M. B., and Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: a review of the literature and a new multi-task investigation. *Acta Psychol.* 115, 43–67.

Gilbert, J. K., Compton, D. L., and Kearns, D. M. (2011). Word and person effects on decoding accuracy: a new look at an old question. *J. Educ. Psychol.* 103, 489-507.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *J. Exp. Psychol. Hum. Percept. Perform.* 5, 674–691.

Goswami, U. (2010). "The relationship between phonological awareness and orthographic representation in different orthographies," in *Learning to Read and Write: A Crosslinguistic Perspective*, eds M. Harris and G. Hatano (Cambridge: Cambridge University Press).

Goswami, U., and Bryant, P. (1990). *Essays in Developmental Psychology Series. Phonological Skills and Learning to Read*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Guan, C. Q., Fraundorf, S. H., and Perfetti, C. A. (2020). Character and child factors contribute to character recognition development among good and poor readers from grade 1 to 6. *Ann. Dyslexia*. doi: 10.1007/s11881-020-00191-0

Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F., and Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *J. Educ. Psychol.* 103, 509–522.

Guan, C. Q., Ye, F., Meng, W. J., and Leong, C. K. (2013). Are poor Chinese text comprehenders also poor in written composition? *Ann. Dyslexia* 63, 217–238. doi: 10.1007/s11881-013-0081-0

Guan, C. Q., Ye, F., Wagner, R. K., Leong, C. K., and Meng, W.-J. (2015). Text comprehension mediates the roles of morphological awareness, syntactic processing, and working memory in predicting Chinese written composition. *J. Educ. Psychol.* 106, 779–798.

Guan, C. Q., Ye, F., Wagner, R. K., and Meng, W. J. (2012). Developmental and individual differences in Chinese writing. *Read. Writ.* 26, 1031–1056.

Guan, C. Q., Zhao, J., Kwok, R. K. W., and Wang, Y. (2019). How does morphosyntactic skill contribute to different genres of Chinese writing from grades 3 to 6? *J. Res. Read.* 42, 239–267.

Gunderson, L., D'Silva, R., and Chen, L. (2011). "Second language reading disability: international themes," in *Handbook of Reading Disability Research*, eds A. McGill-Franzen and R. L. Allington (New York, NY: Routledge), 13–24.

Ho, C. S.-H., Ng, T.-T., and Ng, W.-K. (2003). A "Radical" approach to reading development in Chinese: the role of semantic radicals and phonetic radicals. *J. Lit. Res.* 35, 849–878.

Hohenstein, S., and Kliegl, R. (2020). *Remef: Remove Partial Effects. [R Package Version 1.0.6.9000]*. Available online at: https://github.com/hohenstein/remef/ (accessed April 1, 2020).

Hsu, C.-H., Tsai, J.-L., Lee, C.-Y., and Tzeng, O. J.-L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: an event-related potential study. *Brain Lang.* 108, 56–66. doi: 10.1016/j.bandl.2008.09.002

Hu, C.-F., and Catts, H. W. (1998). The role of phonological processing in early reading ability: What we can learn from Chinese. *Sci. Stud. Read.* 2, 55–79. doi: 10.1111/desc.12793

James, J. S. (1996). *Phonological Processing in Early Reading and Invented Spelling Unpublished Master of Education*. Ph.D. thesis, Queen's University, Canada.

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *J. Mem. Lang.* 46, 723–750.

Johnston, R. A., and Barry, C. (2006). Age of acquisition and lexical processing. *Vis. Cogn.* 13, 789–845.

Kucera, H., and Francis, W. (1967). *Computational Analysis of Present- Day American English*. Providence, RI: Brown University Press.

Kuznetsova, A., Brockhoff, P., and Christensen, H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26.

Laxon, V., Gallagher, A., and Masterson, J. (2002). The effects of familiarity, orthographic neighbourhood density, letter-length and graphemic complexity on children's reading accuracy. *Br. J. Psychol.* 93, 269–287.

Laxon, V. J., Coltheart, V., and Keating, C. (1988). Children find friendly words friendly too: Words with many orthographic neighbours are easier to read and spell. *Br. J. Educ. Psychol.* 58, 103–119.

Lenth, R. (2019). *Emmeans: Estimated Marginal Means, Aka Least-Squares Means [R Package Version 1.3.5]*. Available online at: https://CRAN.R-project.org/package=emmeans (accessed April 1, 2020).

Li, M., Kirby, J. R., Cheng, L., Wade-Woolley, L., and Qiang, H. (2012). Cognitive predictors of English reading achievement in Chinese English-immersion students. *Read. Psychol.* 33, 423–447.

Liberman, I. Y., Shankweiler, D., and Fischer, F. W. (1974). Explicit syllable and phoneme segmentation in the young child. *J. Exp. Child Psychol.* 18, 201–212. doi: 10.1093/deafed/ens067

Marslen-Wilson, W., Tyler, L. K., Walker, R., and Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychol. Rev.* 101, 3–33.

Morrison, C. M., Hirsh, K. W., Chappell, T. D., and Ellis, A. W. (2002). Age and age of acquisition: an evaluation of the cumulative frequency hypothesis. *Eur. J. Cogn. Psychol.* 14, 435–459.

Murray, W. S., and Forster, K. I. (2004). Serial mechanisms in lexical access: the rank hypothesis. *Psychol. Rev.* 111, 721–756.

NIES (2012). Initial Success in Assessment and Intervention on Learning Disabilities at GaoQiao Primary school in Ningbo. Available online at: http://www.nies.net.cn/syq_1353/hzxc/jgcg/201808/t20180822_327435.html (accessed April 1, 2020).

Peng, P., Fuchs, D., Fuchs, L. S., Elleman, A. M., Kearns, D. M., Gilbert, J. K., et al. (2019). A longitudinal analysis of the trajectories and predictors of word reading and reading comprehension development among at-risk readers. *J. Learn. Disabil.* 52, 195–208. doi: 10.1177/0022219418809080

Perfetti, C. A. (1991). "Representations and awareness in the acquisition of reading competence," in *Learning to Read: Basic Research and its Implication*, eds L. Rieben and C. A. Perfetti (Hillsdale, NJ: Erlbaum), 33–44.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Sci. Stud. Read.* 11, 357–383.

Perfetti, C. A., and Harris, L. N. (2013). Universal reading processes are modulated by language and writing system. *Lang. Learn. Dev.* 9, 296–316.

Perfetti, C. A., and Hart, L. (2002). "The lexical quality hypothesis," in *Precursors of functional literacy*, eds L. Verhoeven, C. Elbro, and P. Reitsma (Amsterdam: John Benjamins Company), 189–213.

Perfetti, C. A., and Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *J. Educ. Psychol.* 67, 461–469.

Perfetti, C. A., and Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Sci. Stud. Read.* 18, 22–37.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., and Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol. Rev.* 103, 56–115.

Seidenberg, M. S., and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 4, 523–568.

Seidenberg, M. S., and Waters, G. S. (1985). Spelling-sound effects in reading: time-course and decision criteria. *Mem. Cogn.* 13, 557–572.

Seymour, P. H., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174.

Shu, H., and Anderson, R. C. (1997). Role of orthographic awareness in the character and word acquisition of Chinese children. *Read. Res. Q.* 32, 78–89.

Shu, H., and Anderson, R. C. (1999). "Learning to read Chinese: the development of metalinguistic awareness," in *Reading Chinese script: A cognitive analysis*, eds J. Wang, A. W. Inhoff, and H. C. Chen (Mahwah, NJ: Erlbaum), 1–18.

Shu, H., Chen, X., Anderson, R. C., Wu, N., and Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Dev.* 74, 27–47.

Shu, H., and Wu, N. (2006). "Growth of orthography-phonology knowledge in the Chinese writing system," in *The Handbook of East Asian psycholinguistics*, Vol. 1, eds P. Li, L. H. Tan, E. Bates, and O. J. L. Tzeng (New York, NY: Cambridge University Press), 103–113.

Shu, H., Zhou, X., and Wu, N. (2000). Utilizing phonological cues in Chinese characters: a developmental study. *Acta Psychol. Sin.* 32, 164–169.

Siok, W. T., and Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Dev. Psychol.* 37, 886–899.

Spieler, D. H., and Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychol. Aging* 15, 225–231.

Steacy, L. M., Elleman, A. M., Lovett, M. W., and Compton, D. L. (2016). Exploring differential effects across two decoding treatments on item-level transfer in children with significant word reading difficulties: a new approach for testing intervention elements. *Sci. Stud. Read.* 20, 283–295. doi: 10.1080/10888438.2016.117 8267

Tainturier, M. J., Tremblay, M., and Lecours, A. (1989). Aging and the word frequency effect: a lexical decision investigation. *Neuropsychologia* 27, 1197–1202.

Taraban, R., and McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *J. Mem. Lang.* 26, 608–631.

Treiman, R. (1985). Onsets and rimes as units of spoken syllables: evidence from children. *J. Exp. Child Psychol.* 39, 161–181.

Treiman, R., and Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Dev. Psychol.* 33, 771–780.

Treiman, R., and Zukowski, A. (1991). "Levels of phonological awareness," in *Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman*, eds S. A. Brady and D. P. Shankweiler (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc), 67–83.

Van Orden, G. C., Pennington, B. F., and Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychol. Rev.* 97, 488–522.

Wagner, R. K., and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101, 192–212. doi: 10.1016/j.jecp.2011.11.007

Weekes, B. S., Castles, A. E., and Davies, R. A. (2006). Effects of consistency and age of acquisition on reading and spelling among developing readers. *Read. Writ.* 19, 133–169.

Yang, H., and Peng, D. L. (1997). "The learning and naming of Chinese characters of elementary school children," in *Cognitive Processing of Chinese and Related Asian Languages*, ed. H.-C. Chen (Hong Kong: The Chinese University Press), 323–346.

Zevin, J. D., and Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *J. Mem. Lang.* 47, 1–29. doi: 10.1016/j.cognition.2008.08.004

Visual Dysfunction in Chinese Children With Developmental Dyslexia: Magnocellular-Dorsal Pathway Deficit or Noise Exclusion Deficit?

Yuzhu Ji[1,2] and Hong-Yan Bi[1,2]*

[1] *CAS Key Laboratory of Behavioral Science, Center for Brain Science and Learning Difficulties, Institute of Psychology, Chinese Academy of Sciences, Beijing, China,* [2] *Department of Psychology, University of Chinese Academy of Sciences, Beijing, China*

Many studies have suggested that children with developmental dyslexia (DD) not only show phonological deficit but also have difficulties in visual processing, especially in non-alphabetic languages such as Chinese. However, mechanisms underlying this impairment in vision are still unclear. Visual magnocellular deficit theory suggests that the difficulties in the visual processing of dyslexia are caused by the dysfunction of the magnocellular system. However, some researchers have pointed out that previous studies supporting the magnocellular theory did not control for the role of "noise". The visual processing difficulties of dyslexia might be related to the noise exclusion deficit. The present study aims to examine these two possible explanations via two experiments. In experiment 1, we recruited 26 Chinese children with DD and 26 chronological age–matched controls (CA) from grades 3 to 5. We compared the Gabor contrast sensitivity between the two groups in high-noise and low-noise conditions. Results showed a significant between-group difference in contrast sensitivity in only the high-noise condition. In experiment 2, we recruited another 29 DD and 29 CA and compared the coherent motion/form sensitivity in the high- and low-noise conditions. Results also showed that DD exhibited lower coherent motion and form sensitivities than CA in the high-noise condition, whereas no evidence was observed that the group difference was significant in the low-noise condition. These results suggest that Chinese children with dyslexia have noise exclusion deficit, supporting the noise exclusion hypothesis. The present study provides evidence for revealing the visual dysfunction of dyslexia from the Chinese perspective. The nature of the perceptual noise exclusion and the relationship between the two theoretical hypotheses are discussed.

Keywords: developmental dyslexia, magnocellular theory, noise exclusion, Chinese children, visual dysfunction

INTRODUCTION

The main feature of developmental dyslexia (DD) is a specific and significant impairment in the acquisition of reading skills that is not solely accounted for by mental age, visual acuity problems, or inadequate schooling (World Health Organization, 2011). The phonological deficit theory, which is widely accepted in alphabetic languages, postulates that the difficulties in representation,

storage, or retrieval of speech sounds have a negative impact on the development of grapheme-phenome correspondences, eventually leading to poor phonological skills and reading disability in dyslexia (Snowling, 2001; Ramus, 2003). However, some researchers believe that the specific reading impairments might be traced to some general perceptual processing problems, such as auditory temporal processing impairment (Tallal, 2004), visual magnocellular deficit (Stein, 1997, 2001, 2014), and cerebellar deficit (Nicolson et al., 2001; Nicolson and Fawcett, 2007).

Initially, DD was described as word-blindness, which emphasized the importance of visual processing problems in addition to the phonological deficit. In the late 19th century, studies reported some general visual deficits in dyslexia (Morgan, 1896; Orton, 1925). Later, more and more studies found it was related to the magnocellular pathway deficit. Dyslexics did poorly in processing the rapid visual information that is carried by the visual magnocellular system and the postmortem study also provided evidence that the magnocellular layers of the lateral geniculate nucleus (LGN) in dyslexia was more variable in shape and smaller in general compared with controls (Livingstone et al., 1991; Galaburda and Livingstone, 1993). Therefore, the magnocellular theory was proposed to explain the visual dysfunction of dyslexia (Stein, 1997, 2001). Some researchers also named it as magnocellular-dorsal theory (e.g., Gori et al., 2014), because the dorsal stream mainly received information from the magnocellular pathway (Livingstone and Hubel, 1988; Boden and Giaschi, 2007). In recent decades, many studies found impaired visual magnocellular-dorsal pathway function in dyslexics by means of behavioral and neuroimaging measurements (Boets et al., 2011; Jednoróg et al., 2011), which confirmed the magnocellular theory. However, Sperling et al. (2005, 2006) pointed out that some previous studies that found magnocellular-dorsal deficits in dyslexics used stimuli with noisy conditions, so they assumed that the visual difficulties in DD might be associated with a noise exclusion deficit rather than magnocellular pathway deficit.

Sperling et al. (2005) first used a Gabor contrast sensitivity task to examine their hypothesis. This paradigm is often used to detect the magnocellular pathway function of dyslexia. In their study, the magnocellular and parvocellular stimuli were presented with or without noisy display. They found that children with dyslexia showed lower contrast sensitivity than controls only in the high-noise condition, no matter which type of stimuli were used. After that, Sperling et al. (2006) used the coherent motion task, which is usually used to detect the dorsal pathway function of dyslexia. They measured the coherent motion sensitivities, respectively, in the high-noise condition, in which the contrast of the signal dots was the same as the noise dots, and the low-noise condition, in which the signal dots were red. Results showed that the perceptual threshold of the coherent motion of dyslexics in the high-noise condition was significantly higher than that of controls, whereas the group difference disappeared in the low-noise condition, suggesting the noise exclusion deficit in dyslexia. Some subsequent studies also supported this hypothesis. Northway et al. (2010) used the symbol discrimination task to measure the contrast sensitivity. Results showed that the contrast

sensitivity of DD was lower than that of the control group in the high-noise condition, while no evidence was observed that the group differences were significant in the low-noise condition. Conlon et al. (2012) used the coherent motion task, which included three conditions: (1) low signal contrast with high noise contrast, (2) the same contrasts of signal and noise, and (3) high signal contrast with low noise contrast (Conlon et al., 2012). They found that DD exhibited a higher threshold in conditions except for the low-noise condition.

It seems that deficits in noise exclusion contribute to the etiology of dyslexia, but the studies mentioned above did not take the global form task into account. In previous studies that supported the magnocellular-dorsal theory, the global form task was used as the control condition (non-motion) (Hansen et al., 2001; Conlon et al., 2009). DD showed comparable performance with the control group in this task but exhibited poor coherent motion sensitivity. In the global form task, the contrasts of signal and noise were the same as in the coherent motion condition, which means that the stimuli were also presented in the high-noise condition. If the visual impairments in dyslexia are due to noise rather than motion, it should be observed that DD exhibited poor performance in the high-noise condition not only in the motion task but also in the static task. In addition, in the study by Sperling et al. (2005), the authors did not find the deficit of dyslexia to be specific to the magnocellular stimuli, which was inconsistent with the results of previous studies (e.g., Borsting et al., 1996; Demb et al., 1998; Slaghuis and Ryan, 1999; Kevan and Pammer, 2008). It can be seen that the spatial frequency of Gabor was different in these studies. Early primate studies found that only stimuli with both low spatial frequency [e.g., 1.0 cycles per degree (cpd)] and high temporal frequency (e.g., 10 Hz) were unaffected by the destruction of parvocellular layers but that it induced contrast sensitivity reductions following lesions of magnocellular layers (Merigan and Eskin, 1986; Merigan et al., 1991a,b; Skottun, 2000). It has been proved by functional magnetic resonance imaging studies that the anatomical organization and functional properties of the human LGN showed similar patterns compared with monkey LGN (Schneider et al., 2004; Zhang et al., 2015). However, in Sperling et al.'s (2005), the frequency of magnocellular Gabor was 2 cpd, which might not be completely detected by the magnocellular system.

As compared with alphabetic languages, Chinese as a logographic script has more complex spatial structures without clear grapheme–phoneme corresponding rules. Because of the language specificity, it seems that the deficits of Chinese individuals with dyslexia are different from those with alphabetic dyslexia (Shu et al., 2006; Yang and Bi, 2011; Yang et al., 2013, 2016). Despite the discrepancies, Chinese children with dyslexia also exhibit similar visual processing difficulties. Studies found a lower sensitivity of Chinese DD than typically developing children in the coherent motion task, and the sensitivity was correlated with some reading-related skills such as orthographic awareness, phonological awareness, and picture-naming speed (Meng et al., 2011; Qian and Bi, 2014). Researchers have explained these results as reflecting the magnocellular-dorsal pathway deficit in Chinese children with dyslexia. However, just

as Sperling et al. (2005, 2006) indicated, these studies also used high-noise display, so the results could also be explained as the noise exclusion deficit in Chinese dyslexia. It remains unclear whether the visual dysfunction in Chinese children with dyslexia is attributed to the magnocellular-dorsal deficit or the noise exclusion deficit.

The aim of the present study was to examine the two theoretical hypotheses by two experiments in Chinese children with DD. Experiment 1 used a Gabor contrast sensitivity task in which the magnocellular and parvocellular visual stimuli were presented with high and low external noise. Experiment 2 used a coherent motion task and global form task in the high-noise and low-noise conditions. We hypothesized that if DD showed the magnocellular-dorsal deficit, the worse performance of children with dyslexia should be observed in the M condition of the contrast sensitivity task and coherent motion task; if DD showed the noise exclusion deficit, the worse performance of children with dyslexia should be observed in the high-noise conditions whether the stimuli were related to the M condition/motion or not.

EXPERIMENT 1

Methods

Participants

Fifty-two Chinese children in grades 3–5 were recruited from two primary schools in Beijing. Half of them were DD (14 boys; age range: 8.22–11.95 years) and half were chronological age–matched healthy children (CA; 18 boys; age range: 8.24–11.65 years). The screening criteria of DD were a reading ability test score at least 1.5 standard deviations below grade average in the Standardized Character Recognition Test (Wang and Tao, 1993) and IQ greater than 85 as measured by Raven's Standard Progressive Matrices (Raven et al., 1996). These criteria are widely used in Chinese studies for screening Mandarin-speaking children with dyslexia (e.g., Shu et al., 2006; Wang et al., 2010; Meng et al., 2011; Qian and Bi, 2014). We also measured some reading-related skills for children, including word reading fluency, phonological awareness, morphological awareness, and rapid automatized naming (RAN). Dyslexic children showed significantly worse performance than the controls in all tests except phonological awareness test (marginally significance). It can support the reliability of screening for dyslexia. All participants were right-handed. They had normal hearing

and normal or corrected-to-normal vision without any other neurological abnormalities. This study was approved by the ethics committee of the Institute of Psychology, Chinese Academy of Sciences. Detailed information of each group is shown in **Table 1**.

Measures of Reading-Related Skills

Standardized character recognition test

In this test, participants were instructed to write down a compound word with each of the target morpheme characters. Characters are divided into 10 groups based on reading difficulty (206 characters for 3rd graders, 174 characters for 4th graders, and 210 characters for 5th graders). Each correct response was given one point. The score for each group of characters was calculated by multiplying the total points by the corresponding coefficient of difficulty. The final score for each participant was the sum of sub-scores for all 10 character groups to estimate of the number of Chinese characters the children actually recognize.

Word reading fluency

This task contains 160 single Chinese characters with high frequency. Children were asked to read all this words as fast as possible in 1 min. The number of the correct answer were the final score of the task.

Phonological awareness

In this task, children were orally presented with three syllables of Chinese characters and were asked to judge which syllable was different from the others in initial consonant, vowels or tone (e.g., /meng3/was different from/gao1/and/bao4/in vowels). There were 30 items in total and the final score was the number of correct items.

Morphological awareness

In this task, children were presented with one pair of 2-morpheme words which contains the same morpheme (e.g., "老家" rat and "老家" home). They need to judge if the same morpheme in different words has the same meaning. There were 20 items in total and the final score was the number of correct items.

Rapid automatized naming (RAN)

Performance of children's RAN of pictures and digits were collected. Five pictures (flower, book, dog, hand, and shoes) and five digits (2, 4, 6, 7, and 9) were used, respectively, in the two tasks. Pictures/digits were repeatedly presented visually in random order on a 6 × 5 row-column grid. Participants were asked to name each picture/digit in sequence as quickly

TABLE 1 | Characteristics of the two groups in experiment 1 ($M \pm SD$).

	CA ($n = 26$)	DD ($n = 26$)	t	p	Cohen's d
Age (years)	10.25 ± 0.95	10.15 ± 1.05	0.35	0.73	0.10
IQ (standard score)	112.62 ± 12.18	111.04 ± 12.05	0.47	0.64	0.13
Reading ability (number of Chinese characters)	2870.48 ± 540.78	1875.45 ± 531.73	6.70	<0.001	1.86
Word reading fluency (number of correct words/1 min)	107.31 ± 18.92	76.69 ± 17.24	6.10	<0.001	1.69
Phonological awareness (number of correct items)	19.65 ± 5.84	16.73 ± 6.13	1.76	0.085	0.49
Morphological awareness (number of correct items)	25.27 ± 5.19	21.00 ± 5.93	2.76	0.008	0.77
RAN (sec)	14.76 ± 3.05	16.37 ± 2.47	−2.10	0.04	−0.58

FIGURE 1 | Examples of M and P Gabor with high noise and low noise.
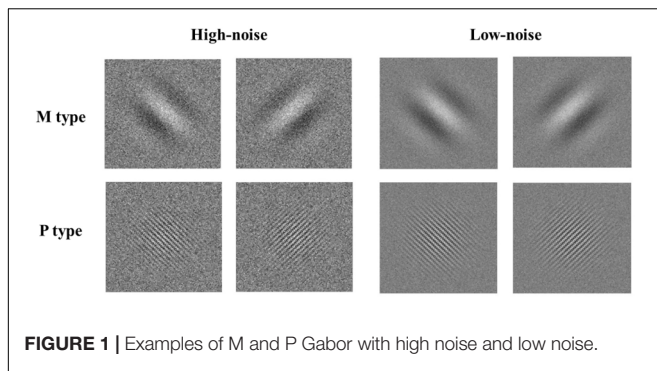
as possible. The total naming time was collected. Each task was conducted twice, and the average score was used as the final RAN score.

Stimuli and Procedure

As shown in **Figure 1**, stimuli consisted of a Gabor pattern of sine wave gratings with checkerboard noise. The magnocellular-type gratings had a spatial frequency of 0.5 cpd and flickered in counter phase at a rate of 15 reversals/s (temporal frequency = 15 Hz). The parvocellular type gratings had a spatial frequency of 5 cpd and did not reverse phase (temporal frequency = 5 Hz). Both of these kinds of gratings had two orientations (45° or 135°). Noise consisted of 2 × 2 pixel patches. The contrast of each pixel patch was sampled from a Gaussian distribution. In the high-noise condition, the contrast of the brightest and darkest pixel patches was 100%; in the low-noise condition, it was 40%, selected by a pilot study to avoid the ceiling effects. Noise also reversed phase when accompanied by M stimuli but was static when accompanied by P stimuli. The size of each kind of stimuli was 6° × 6°.

The task was programmed using Matlab R2015b with Psychtoolbox extensions. The monitor resolution was 1366 × 768, and its vertical refresh rate was 60 Hz. Stimuli were shown on a gray background with a luminance of 51.73 cd/m². Children sat 60 cm from the computer screen and were given the opportunity to practice. In the formal experiment, a fixation was first shown at the center of the screen for 250 ms, and then the stimuli appeared. After 200 ms, a blank screen was presented, and children were asked to judge the orientation of stimuli by pressing the corresponding keys without time limitation. The contrast of Gabor in a single trail was determined by a 3-down/1-up staircase. The initial contrast was 50%. Before the first reversal, a step amounted to change contrast by 20% of the present contrast level. After that, it was changed by 10% of the present level. The program stopped when children reached 150 trials or 10 times of reversal. The average contrast level for the last five reversals was taken to estimate the contrast threshold. Four separated staircases were applied for different conditions, and the order was counterbalanced across participants.

Data Analysis

The three-way repeated-measures analysis of variance (ANOVA) was conducted firstly, with a between-subject factor (group:

DD, CA) and two within-subject factors (stimulus type: M, P; noisy condition: high noise, low noise). Then the two-way repeated-measures ANOVAs were conducted for the high-noise condition and low-noise condition, respectively, with a between-subject factor (group: DD, CA) and a within-subject factor (stimuli type: M, P).

Results

The thresholds of M/P Gabor with high noise and low noise in the two groups are presented in **Table 2**.

It showed that the three-way interaction was marginally significant ($F_{1,50} = 3.62$, $p = 0.063$, partial $\eta^2 = 0.068$). In order to better understand this effect, the two-way ANOVAs were further conducted for the high-noise and low-noise conditions separately.

High-Noise Condition

There was a significant main effect of stimuli type ($F_{1,50} = 5.20$, $p = 0.027$, partial $\eta^2 = 0.094$), with a higher threshold for P stimuli than M stimuli. The main effect of group was also significant ($F_{1,50} = 6.14$, $p = 0.017$, partial $\eta^2 = 0.109$). DD exhibited a higher threshold than CA. The interaction between group and stimuli type was non-significant ($F_{1,50} = 1.99$, $p = 0.165$, partial $\eta^2 = 0.038$).

Low-Noise Condition

The ANOVA showed non-significant main effects of group ($F_{1,50} = 2.30$, $p = 0.136$, partial $\eta^2 = 0.044$) and stimuli type ($F_{1,50} = 0.36$, $p = 0.553$, partial $\eta^2 = 0.007$). The interaction between group and stimuli type was also non-significant ($F_{1,50} = 0.26$, $p = 0.614$, partial $\eta^2 = 0.005$). See **Figure 2**.

Discussion

In experiment 1, we used the Gabor contrast sensitivity task to investigate the magnocellular/parvocellular pathway function and the role of noise in Chinese children with dyslexia. Results showed that, in only the high-noise condition, dyslexia exhibited significantly lower sensitivities than the control group no matter what type of stimuli they processed. In the low-noise condition, none of the main effects and no interaction was found. These results indicated that Chinese children with dyslexia had noise exclusion deficit, supporting the noise exclusion hypothesis.

Even though more strict parameters were used to set up the magnocellular and parvocellular Gabor as compared with Sperling et al.'s (2005), we still did not find the selective deficit of dyslexia in processing M-type stimuli. This was not in

TABLE 2 | Contrast thresholds (%) for different stimuli in DD and CA ($M \pm SD$).

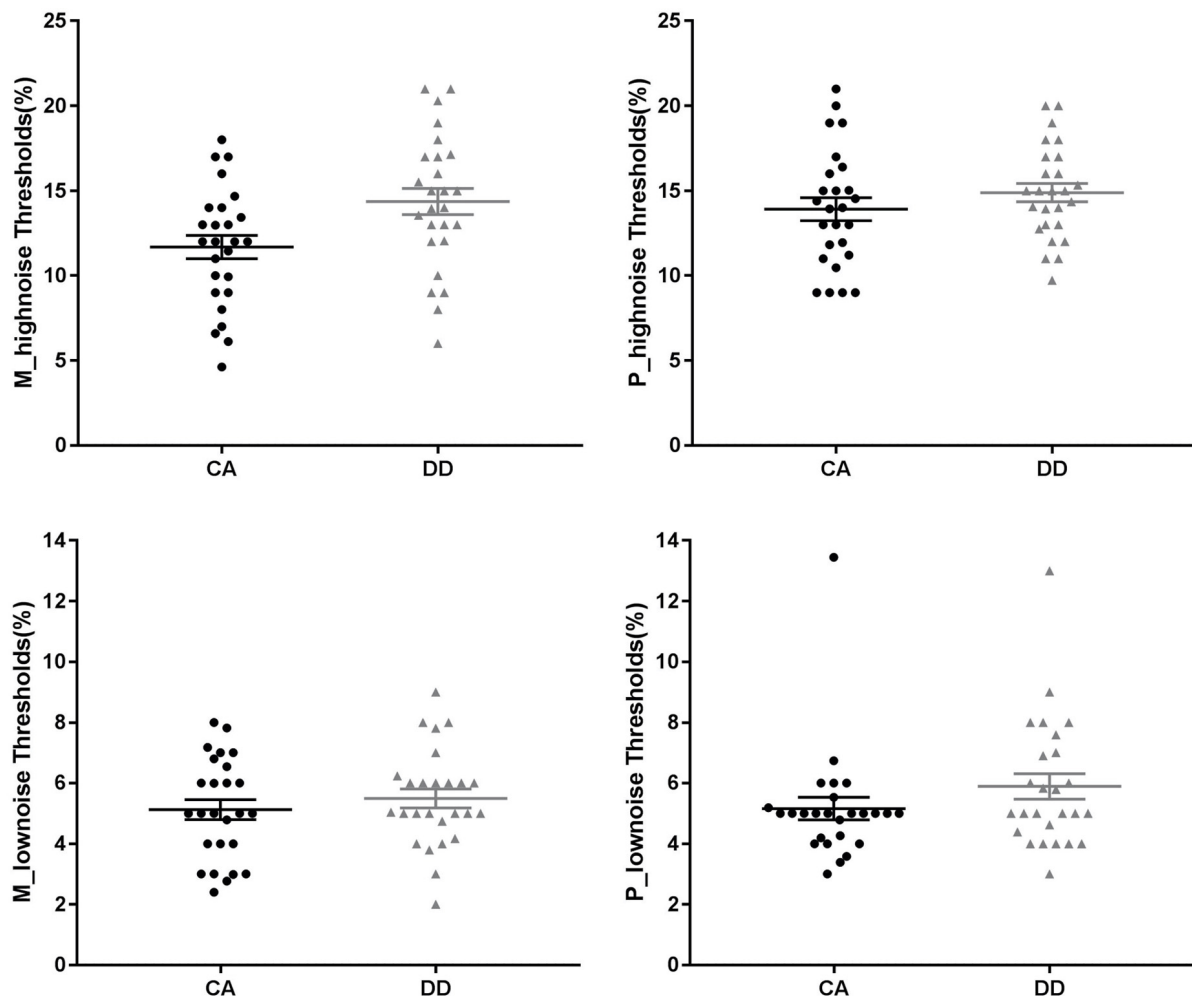|	High-noise		Low-noise	
	M type	**P type**	**M type**	**P type**
CA	11.68 ± 3.50	13.91 ± 3.45	5.1 ± 1.7	5.2 ± 1.9
DD	14.36 ± 3.95	14.89 ± 2.74	5.5 ± 1.6	5.9 ± 2.1

FIGURE 2 | M/P-type Gabor contrast thresholds of two groups in the different noisy conditions. The longest line in the middle denotes the means and the other two lines denote the standard error. CA, chronological age–matched controls; DD, developmental dyslexia.

line with expectations and inconsistent with previous studies (Borsting et al., 1996; Demb et al., 1998; Slaghuis and Ryan, 1999; Kevan and Pammer, 2008). The reasons might be as follows. First, this task involved only the spatial frequency and temporal frequency of Gabor to discriminate M-type and P-type stimuli. The information about contrast and color were not taken into account. Actually, magnocellular layers not only preferred higher temporal frequency and lower spatial frequency but were also sensitive to lower contrast and color-blindness; parvocellular layers preferred lower temporal frequency, higher spatial frequency, and higher contrast and showed robust response to both chromatic and achromatic stimuli (Zhang et al., 2015). Second, the neuronal responses in magnocellular layers and parvocellular layers were preferentially rather than exclusively tuned to M-type and P-type stimuli (Skottun, 2015; Zhang et al., 2016). A behavioral experiment cannot directly measure the responses of M layers and P layers to different types of stimuli; thus, it might not be able to detect such a subtle deficit.

EXPERIMENT 2
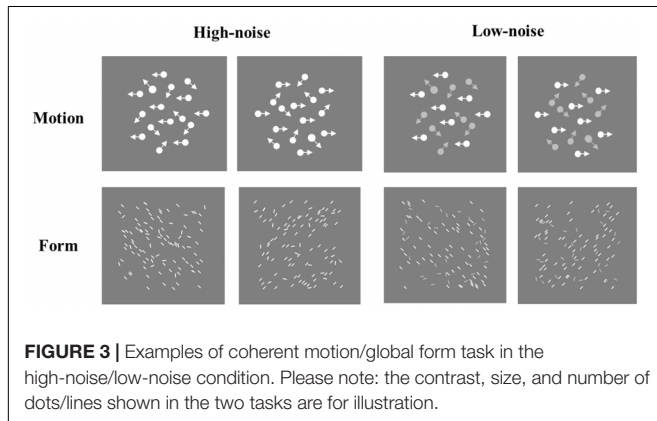
Methods

Participants

Another 58 Chinese children in grades 3–5 were recruited. Half were DD (21 boys; age range: 8.78–11.51 years) and half were CA (22 boys; age range: 8.63–11.45 years). All participants were right-handed. The screening criteria of DD and CA were same as in experiment 1. We also measured some reading-related skills for children (same as experiment 1). Dyslexic children showed significantly worse performance than the controls in all tests. It can support the reliability of screening for dyslexia. Detailed information about each group is shown in **Table 3**.

Stimuli and Procedure

The coherent motion stimuli were generated by a random-dot kinematogram, which comprised 100 moving white dots with a diameter of 0.14° and a speed of 2°/s. The signal dots moved

TABLE 3 | Characteristics of the two groups in experiment 2 ($M \pm SD$).

	CA (n = 29)	DD (n = 29)	t	p	Cohen's d
Age (years)	10.50 ± 0.85	10.36 ± 0.92	0.60	0.55	0.16
IQ (standard score)	112.07 ± 12.18	108.62 ± 13.43	1.03	0.31	0.27
Reading ability (number of Chinese characters)	2881.64 ± 261.72	1984.26 ± 325.37	11.57	<0.001	3.04
Word reading fluency (number of correct words/1 min)	105.03 ± 18.38	74.24 ± 13.27	7.32	<0.001	1.92
Phonological awareness (number of correct items)	27.24 ± 2.82	24.90 ± 4.84	2.26	0.03	0.59
Morphological awareness (number of correct items)	16.83 ± 1.73	14.93 ± 2.05	3.80	<0.001	1.00
RAN (sec)	13.82 ± 1.88	15.89 ± 2.04	−4.00	<0.001	−1.05

FIGURE 3 | Examples of coherent motion/global form task in the high-noise/low-noise condition. Please note: the contrast, size, and number of dots/lines shown in the two tasks are for illustration.

coherently in a single direction (left or right), and the noise dots moved randomly. To prevent eye tracking, each dot had a lifetime of 3 frames, after which the dot disappeared and was regenerated at a randomly selected location (the radius of moving scope ranged from 1° to 4°). Compared with the coherent motion task, we designed a new global form task, which comprised 100 static lines in a 6° × 6° area. The size of each line was 0.26° × 0.06°. The orientation of signal lines was fixed (45° or 135°) and that of noise lines was random. Both tasks had two versions: high noise and low noise, respectively. In the high-noise condition, the signal contrast was the same as the noise contrast (both were 63.88%); in the low-noise version, the signal contrast was also 63.88%, but the noise contrast was 58.52% (see **Figure 3**). These contrasts were selected by a pilot study to avoid the ceiling effects in the low-noise condition.

All tasks were also programmed by Matlab R2015b with Psychtoolbox extensions. The monitor resolution was 1366 × 768, and its vertical refresh rate was 60 Hz. Stimuli were shown on a gray background with luminance of 12.98 cd/m². Children sat 60 cm from the computer screen and were given the opportunity to practice. The procedure of the two tasks was quite similar. A fixation was first shown at the center of the screen for 250 ms. Then it disappeared in the global form task but remained on screen through one single trial in the coherent motion task. Stimuli were shown for 1000 ms. After that, a blank screen was presented, and children were asked to judge the direction of the signal dots or the orientation of signal lines by pressing the corresponding keys with no time limitation. The proportion of signals in a single trail was determined by a

3-down/1-up staircase. The initial proportion was 50%. Before the first reversal, a step amounted to the change proportion of signals by 20% of the present proportion level. After that, it was changed by 10% of the present level. The program stopped when children reached 150 trials or 10 times of reversal. The average proportion level for the last five reversals was used to estimate the threshold. Four separated staircases were applied for the two versions of two tasks, and the order was counterbalanced across participants.

Data Analysis

The three-way repeated-measures ANOVA was conducted firstly, with a between-subject factor (group: DD, CA) and two within-subject factors (task: motion, form; noisy condition: high noise, low noise). Then the two-way repeated-measures ANOVAs were conducted for the high-noise condition and low-noise condition, respectively, with a between-subject factor (group: DD, CA) and a within-subject factor (task: motion, form).

Results

The thresholds of coherent motion and global form of the two groups in the high-noise and low-noise conditions are shown in **Table 4**.

It showed that the three-way interaction was non-significant ($F_{1,56} = 0.69$, $p = 0.410$, partial $\eta^2 = 0.012$). The interaction between noise and group did not reach significance ($F_{1,56} = 0.99$, $p = 0.323$, partial $\eta^2 = 0.017$), but the interaction between task and noise was significant ($F_{1,56} = 4.29$, $p = 0.043$, partial $\eta^2 = 0.071$). We proceed with the two-way ANOVAs separately for the high-noise and low-noise conditions, in order to further understand the results and keep consistent with experiment 1.

High-Noise Condition

The ANOVA showed a significant main effect of group ($F_{1,56} = 5.96$, $p = 0.018$, partial $\eta^2 = 0.096$), and DD exhibited a higher threshold than CA. Neither the main effect of task

TABLE 4 | Thresholds (%) for different tasks and conditions in DD and CA ($M \pm SD$).

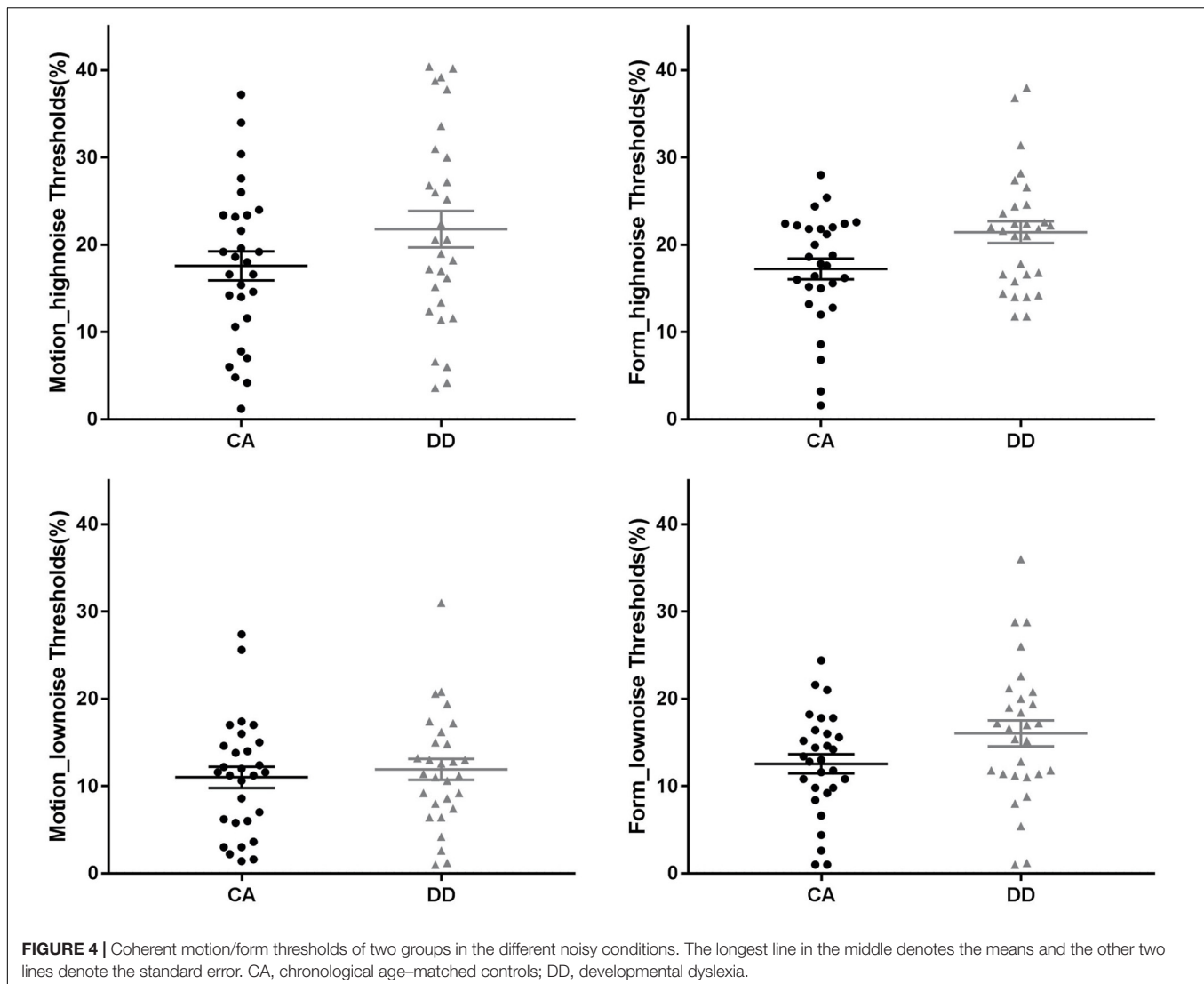|	High-noise		Low-noise	
	Motion	**Form**	**Motion**	**Form**
CA	17.59 ± 8.93	17.23 ± 6.37	11.00 ± 6.57	12.56 ± 5.84
DD	21.79 ± 11.25	21.44 ± 6.70	11.91 ± 6.49	16.05 ± 7.97

FIGURE 4 | Coherent motion/form thresholds of two groups in the different noisy conditions. The longest line in the middle denotes the means and the other two lines denote the standard error. CA, chronological age–matched controls; DD, developmental dyslexia.

($F_{1,56}$ = 0.06, p = 0.807, partial η^2 = 0.001) nor the interaction between group and task was significant ($F_{1,56}$ < 0.001, p = 0.996, partial η^2 < 0.001).

Low-Noise Condition
The main effect of task was significant ($F_{1,56}$ = 9.40, p = 0.003, partial η^2 = 0.144), in that the threshold for the form task was higher than that for the motion task. However, neither the main effects of group ($F_{1,56}$ = 2.11, p = 0.152, partial η^2 = 0.036) and task nor the interaction ($F_{1,56}$ = 1.93, p = 0.171, partial η^2 = 0.033) was significant. See **Figure 4**.

Discussion
Results of experiment 2 showed that DD exhibited a higher threshold than CA in the high-noise condition, whereas no evidence was observed that the group difference was significant in the low-noise condition. This suggests that Chinese children with dyslexia have noise exclusion deficit, whether it is related to motion or not, also supporting the noise exclusion hypothesis.

One of the main findings in experiment 2 was that Chinese children with dyslexia showed a noise exclusion deficit in the coherent motion task, which is the same as the results from alphabetic languages studies (Sperling et al., 2006; Conlon et al., 2009; Northway et al., 2010). This revealed that the visual difficulties of Chinese DD were related to noise rather than motion, and the noise exclusion deficit in DD might be a cultural-general deficit. In addition, results of experiment 2 also showed a higher threshold of dyslexic children in the global form task with high-noise condition. This result was inconsistent with the previous studies of Hansen et al. (2001), Conlon et al. (2009), and Meng et al. (2011). In their studies, stimuli were presented only in the high-noise condition, and the poor sensitivity of dyslexia was observed in the coherent motion task rather than global form task. The possible reason for the inconsistent results might be the different difficulties of the two tasks. Thresholds of the global form task were higher than that of the coherent motion task observed in those three studies. This might result in a possible floor effect in the global form

task, resulting in the inability to find a difference between the two groups. In the present study, no evidence was observed that the task main effect was significant in the high-noise condition, which means that the difficulties of the two tasks were the same. In this case, we found only a significant group main effect, suggesting that Chinese children with dyslexia have poor coherent sensitivities and that this is related to the noise rather than stimuli type.

GENERAL DISCUSSION

The present study examined two theoretical hypotheses to explain the visual dysfunction of Chinese children with dyslexia. Two experiments consistently showed that dyslexic children showed poorer performance than controls only in the high-noise condition no matter what kind of stimuli types and tasks they processed. This suggests that Chinese children with dyslexia have a noise exclusion deficit, supporting the noise exclusion hypothesis. The present study provides evidence for revealing the cognitive mechanism of visual dysfunction in dyslexia from the Chinese perspective.

Noise Exclusion Deficit in Chinese Children With Dyslexia

This study was based on the two previous studies of Sperling et al. (2005, 2006) and improved the experimental paradigm. In experiment 1, we used strict spatial frequency and temporal frequency for the M Gabor and P Gabor. In experiment 2, we designed a new global form task as a control to the motion task. In that case, the results of the two experiments still showed the noise exclusion deficit of Chinese children with dyslexia, which was consistent with the results of Sperling et al.'s studies in an alphabetic language cultural context. This might suggest that dyslexia have a relatively robust noise exclusion deficit across different language cultures. Despite the discrepancies between different language systems, Chinese children with dyslexia also exhibited the same cognitive mechanism of their visual processing difficulties. Similarly, a previous neuroimaging study also found a common brain activation for semantic decisions on written words in Chinese and English dyslexics despite different activation in Chinese versus English normal readers (Hu et al., 2010).

Given that Chinese children with dyslexia showed noise exclusion deficit, how might it affect reading acquisition? Sperling et al. (2005) proposed three possibilities. (1) The visual impairment is part of a broader problem with noise exclusion that affects speech and further influence reading. (2) The deficit directly affects reading through the visual modality. (3) The visual deficit could have detrimental effects on the development of phonological representations and then affect reading acquisition. For Chinese reading, the effects of noise exclusion deficit on reading impairments might be the possibilities of (2) and (3). First, no matter in Chinese or alphabetic language reading, word recognition requires abstracting away from variations in size, font, and style. It

may be more difficult if visual processing is hampered by deficits in noise exclusion. Sperling et al. (2005) second, although different from the letter-by-letter phonemic segments in alphabetic languages, the experience with phonetic radicals of Chinese characters also shapes the development of phonological information. If children have difficulties in extracting phonetic information from noisy distractors, phonological presentation would be affected.

Noise Exclusion and Visual-Spatial Attention

Given that the noise exclusion deficit might be a cross-cultural deficit, what is the nature of it? In the present study, results showed that dyslexic children exhibited poor contrast sensitivities and coherent sensitivities than controls only in the high-noise condition. It might reveal that for children with dyslexia, the distractors were more difficult to inhibit. Researchers proposed that signal enhancement and noise exclusion (inhibition of distractors) are two mechanisms of visual-spatial attention to optimize perceptual judgment. Noise exclusion can help to improve the perceptual filtering so that signals are processed and noise is excluded (Sperling et al., 2006). However, the invalid attention window of DD during processing will expose the target stimuli to the spatial noisy distractors (Facoetti et al., 2008). Therefore, some researchers believe that the noise exclusion deficit shown in DD is essentially caused by visual-spatial attention deficit (Facoetti et al., 2008). Previous studies have shown the attention impairments of dyslexia: individuals with dyslexia cannot shift their attention from one window to another and have a prolonged attentional dwell time, suggesting their sluggish attentional shifting (Hari and Renvall, 2001). Effective attention shifting plays an important role in reading. However, people with dyslexia exhibited spatial sluggish attentional shifting (e.g., Ruffino et al., 2010b, 2014; Vidyasagar and Pammer, 2010) in visual sense modalities (Facoetti et al., 2010), which might finally lead to the poor reading performance. Some studies found that visual selective attention deficits in dyslexia may be due to a specific difficulty in orienting and focusing and a diffused distribution of visual processing resources (Facoetti et al., 2000a,b). Other studies also found that the noise exclusion deficit in DD could be moderated by visual-spatial attention (Ruffino et al., 2010a; Conlon et al., 2012).

The Relationship Between Noise Exclusion Hypothesis and Magnocellular Theory

Even though the findings of this study supported noise exclusion hypothesis, it still cannot exclusively rule out magnocellular theory. The magnocellular theory was a neural physiological interpretation of visual deficits in dyslexia, but the noise exclusion hypothesis was described as a behavioral level theory. In actuality, the noise exclusion might also have its underlying neural mechanism. As mentioned above, it is undeniable that

perceptual noise exclusion is closely related to visual-spatial attention. In the frontoparietal network, the parietal posterior cortex (PPC) is one of the essential areas for visual spatial attention (Saalmann et al., 2007). It seems that the noise exclusion is related to the function of PPC. In addition, PPC is also considered as a part of the magnocellular-dorsal pathway (Saalmann et al., 2007). Therefore, researchers argued that noise exclusion might be related to the magnocellular-dorsal pathway. Because of the large visual receptive field and the fast conduction velocity, the magnocellular system provides an initial rapid, low-spatial-frequency signal, possibly through the dorsal stream to the parietal and frontal regions (Vidyasagar, 2005). This early activation is thought to provide an initial global analysis of the object foreground/background segregation, before feedback signals into the inferotemporal cortex fill in the details (Laycock, 2012). This indicates that the magnocellular-dorsal pathway theory and the noise exclusion hypothesis are not two completely opposite theoretical hypothesis, especially in the brain network. We believe they may reflect DD's dysfunction in different levels of the visual system. The magnocellular-dorsal theory may emphasize the atypical functional characteristics of different stages of the visual conduction pathway in dyslexia, especially the early stages, while the noise exclusion hypothesis may emphasize the abnormal top-down regulation by the high-order cortex of early visual processing in dyslexia. This hypothesis should be examined in future research using a brain imaging method.

LIMITATIONS

There were several limitations to this study. First, we only used a "low achievement" criterion to screen the Chinese children with dyslexia (reading ability score below -1.5 SD), but it was not a mainstream in an international context. We should use the "persistence" and/or "resistance" criterion to screen dyslexia strictly in the future. Second, the participants in two experiments were not the same group of children. We will further test the reliability of the results in the same group of children in the future. Third, in experiment 1, the Gabor contrast sensitivity task manipulated the contrast ratio of stimuli. It cannot probe the function of M or P pathway strictly, because the M pathway was sensitive to the stimuli with high temporal frequency, low spatial frequency and low contrast, and the P pathway was sensitive to the stimuli with low temporal frequency, high spatial frequency and high contrast. Future research should design a better paradigm to more strictly detect the function of M and P pathway.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institute of Psychology, Chinese Academy of Sciences. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YJ and H-YB designed the study. YJ conducted the experiments, analyzed the data, wrote and revised the manuscript with the help of H-YB.

FUNDING

REFERENCES

Boden, C., and Giaschi, D. (2007). M-stream deficits and reading-related visual processes in developmental dyslexia. *Psychol. Bull.* 133, 346–366. doi: 10.1037/0033-2909.133.2.346

Boets, B., Vandermosten, M., Cornelissen, P., Wouters, J., and Ghesquière, P. (2011). Coherent motion sensitivity and reading development in the transition from prereading to reading stage. *Child Dev.* 82, 854–869. doi: 10.1111/j.1467-8624.2010.01527.x

Borsting, E., Ridder, W. H., Dudeck, K., Kelley, C., Matsui, L., and Motoyama, J. (1996). The presence of a magnocellular defect depends on the type of dyslexia. *Vis. Res.* 36, 1047–1053. doi: 10.1016/0042-6989(95)00199-9

Conlon, E. G., Lilleskaret, G., Wright, C. M., and Power, G. F. (2012). The influence of contrast on coherent motion processing in dyslexia. *Neuropsychologia* 50, 1672–1681. doi: 10.1016/j.neuropsychologia.2012.03.023

Conlon, E. G., Sanders, M. A., and Wright, C. M. (2009). Relationships between global motion and global form processing, practice, cognitive and visual processing in adults with dyslexia or visual discomfort. *Neuropsychologia* 47, 907–915. doi: 10.1016/j.neuropsychologia.2008.12.037

Demb, J. B., Boynton, G. M., and Heeger, D. J. (1998). Functional magnetic resonance imaging of early visual pathways in dyslexia. *J. Neurosci.* 18, 6939–6951. doi: 10.1523/JNEUROSCI.18-17-06939.1998

Facoetti, A., Paganoni, P., and Lorusso, M. L. (2000a). The spatial distribution of visual attention in developmental dyslexia. *Exp. Brain Res.* 132, 531–538. doi: 10.1007/s002219900330

Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., and Mascetti, G. G. (2000b). Visual-spatial attention in developmental dyslexia. *Cortex* 36, 109–123. doi: 10.1016/S0010-9452(08)70840-2

Facoetti, A., Ruffino, M., Peru, A., Paganoni, P., and Chelazzi, L. (2008). Sluggish engagement and disengagement of non-spatial attention in dyslexic children. *Cortex* 44, 1221–1233. doi: 10.1016/j.cortex.2007.10.007

Facoetti, A., Trussardi, A. N., Ruffino, M., Lorusso, M. L., Cattaneo, C., Galli, R., et al. (2010). Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *J. Cogn. Neurosci.* 22, 1011–1025. doi: 10.1162/jocn.2009.21232

Galaburda, A., and Livingstone, M. (1993). Evidence for a magnocellular defect in developmental dyslexia. *Ann. N. Y. Acad. Sci.* 682, 70–82. doi: 10.1111/j.1749-6632.1993.tb22960.x

Gori, S., Cecchini, P., Molteni, M., and Facoetti, A. (2014). Magnocellular-dorsal pathway and sub-lexical route in developmental dyslexia. *Front. Hum. Neurosci.* 8:460. doi: 10.3389/fnhum.2014.00460

Hansen, P. C., Stein, J. F., Orde, S. R., Winter, J. L., and Talcott, J. B. (2001). Are dyslexics' visual deficits limited to measures of dorsal stream function? *Neuroreport* 12, 1527–1530. doi: 10.1097/00001756-200105250-00045

Hari, R., and Renvall, H. (2001). Impaired processing of rapid stimulus sequences in dyslexia. *Trends Cogn. Sci.* 5, 525–532. doi: 10.1016/S1364-6613(00)01801-5

Hu, W., Lee, H. L., Zhang, Q., Liu, T., Geng, L. B., Seghier, M. L., et al. (2010). Developmental dyslexia in Chinese and English populations: dissociating the effect of dyslexia from language differences. *Brain* 133, 1694–1706. doi: 10.1093/brain/awq106

Jednoróg, K., Marchewka, A., Tacikowski, P., Heim, S., and Grabowska, A. (2011). Electrophysiological evidence for the magnocellular-dorsal pathway deficit in dyslexia. *Dev. Sci.* 14, 873–880. doi: 10.1111/j.1467-7687.2011.01037.x

Kevan, A., and Pammer, K. (2008). Making the link between dorsal stream sensitivity and reading. *NeuroReport* 19, 467–470. doi: 10.1097/WNR.0b013e3282f5f7ad

Laycock, R. (2012). Abrupt and ramped flicker-defined form shows evidence for a large magnocellular impairment in dyslexia. *Neuropsychologia* 50, 2107–2113. doi: 10.1016/j.neuropsychologia.2012.05.012

Livingstone, M., and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240, 740–749. doi: 10.1126/science.3283936

Livingstone, M. S., Rosen, G. D., Drislane, F. W., and Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 88, 7943–7947. doi: 10.1073/pnas.90.6.2556e

Meng, X., Cheng-Lai, A., Zeng, B., Stein, J. F., and Zhou, X. (2011). Dynamic visual perception and reading development in Chinese school children. *Ann. Dyslexia* 61, 161–176. doi: 10.1007/s11881-010-0049-2

Merigan, W. H., and Eskin, T. A. (1986). Spatio-temporal vision of macaques with severe loss of P beta retinal ganglion cells. *Vis. Res.* 26, 1751–1761. doi: 10.1016/0042-6989(86)90125-2

Merigan, W. H., Byrne, C. E., and Maunsell, J. H. (1991a). Does primate motion perception depend on the magnocellular pathway? *J. Neurosci.* 11, 3422–3429. doi: 10.1523/JNEUROSCI.11-11-03422.1991

Merigan, W. H., Katz, L. M., and Maunsell, J. H. (1991b). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *J. Neurosci.* 11, 994–1001. doi: 10.1523/JNEUROSCI.11-04-00994.1991

Morgan, W. P. (1896). A case of congenital word blindness. *Br. Med. J.* 2, 1378.

Nicolson, R. I., and Fawcett, A. J. (2007). Procedural learning difficulties: reuniting the developmental disorders? *Trends Neurosci.* 30, 135–141. doi: 10.1016/j.tins.2007.02.003

Nicolson, R. I., Fawcett, A. J., and Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends Neurosci.* 24, 508–511. doi: 10.1016/S0166-2236(00)01896-8

Northway, N., Manahilov, V., and Simpson, W. (2010). Coloured filters improve exclusion of perceptual noise in visually symptomatic dyslexics. *J. Res. Read.* 33, 223–230. doi: 10.1111/j.1467-9817.2009.01409.x

Orton, S. T. (1925). Word-blindness in school children. *Arch. Neurol. Psychiatry* 14, 581–615.

Qian, Y., and Bi, H.-Y. (2014). The visual magnocellular deficit in Chinese-speaking children with developmental dyslexia. *Front. Psychol.* 5:692. doi: 10.3389/fpsyg.2014.00692

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Curr. Opin. Neurobiol.* 13, 212–218. doi: 10.1016/S0959-4388(03)00035-7

Raven, J. C., Court, J. H., and Raven, J. (1996). *Standard Progressive Matrices*. Oxford: Oxford Psychologists Press.

Ruffino, M., Gori, S., Boccardi, D., Molteni, M., and Facoetti, A. (2014). Spatial and temporal attention in developmental dyslexia. *Front. Hum. Neurosci.* 8:331. doi: 10.3389/fnhum.2014.00331

Ruffino, M., Gori, S., Franceschini, S., and Facoetti, A. (2010a). Developmental dyslexia: perceptual noise exclusion deficit or spatial attention dysfunction? *Perception* 39:80.

Ruffino, M., Trussardi, A. N., Gori, S., Finzi, A., Giovagnoli, S., Menghini, D., et al. (2010b). Attentional engagement deficits in dyslexic children. *Neuropsychologia* 48, 3793–3801. doi: 10.1016/j.neuropsychologia.2010.09.002

Saalmann, Y. B., Pigarev, I. N., and Vidyasagar, T. R. (2007). Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. *Science* 316, 1612–1615. doi: 10.1126/science.1139140

Schneider, K. A., Richter, M. C., and Kastner, S. (2004). Retinotopic organization and functional subdivisions of the human lateral geniculate nucleus: a high-resolution functional magnetic resonance imaging study. *J. Neurosci.* 24, 8975–8985. doi: 10.1523/JNEUROSCI.2413-04.2004

Shu, H., McBride-Chang, C., Wu, S., and Liu, H. (2006). Understanding chinese developmental dyslexia: morphological awareness as a core cognitive construct. *J. Educ. Psychol.* 98, 122–133. doi: 10.1037/0022-0663.98.1.122

Skottun, B. C. (2000). The magnocellular deficit theory of dyslexia: the evidence from contrast sensitivity. *Vis. Res.* 40, 111–127. doi: 10.1016/S0042-6989(99)00170-4

Skottun, B. C. (2015). On the use of spatial frequency to isolate contributions from the magnocellular and parvocellular systems and the dorsal and ventral cortical streams. *Neurosci. Biobehav. Rev.* 56, 266–275. doi: 10.1016/j.neubiorev.2015.07.002

Slaghuis, W. L., and Ryan, J. F. (1999). Spatio-temporal contrast sensitivity, coherent motion, and visible persistence in developmental dyslexia. *Vis. Res.* 39, 651–668. doi: 10.1016/S0042-6989(98)00151-5

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia* 7, 37–46. doi: 10.1002/dys.185

Sperling, A. J., Lu, Z.-L., Manis, F. R., and Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nat. Neurosci.* 8, 862–863. doi: 10.1038/nn1474

Sperling, A. J., Lu, Z.-L., Manis, F. R., and Seidenberg, M. S. (2006). Motion-perception deficits and reading impairment: it's the noise, not the motion. *Psychol. Sci.* 17, 1047–1053. doi: 10.1111/j.1467-9280.2006.01825.x

Stein, J. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends Neurosci.* 20, 147–152. doi: 10.1016/S0166-2236(96)01005-3

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia* 7, 12–36. doi: 10.1002/dys.186

Stein, J. (2014). Dyslexia: the role of vision and visual attention. *Curr. Dev. Disord. Rep.* 1, 267–280. doi: 10.1007/s40474-014-0030-6

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nat. Rev. Neurosci.* 5, 721–728. doi: 10.1038/nrn1499

Vidyasagar, T. R. (2005). attentional gating in primary visual cortex: a physiological basis for dyslexia. *Perception* 34, 903–911. doi: 10.1068/p5332

Vidyasagar, T. R., and Pammer, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends Cogn. Sci.* 14, 57–63. doi: 10.1016/j.tics.2009.12.003

Wang, J.-J., Bi, H.-Y., Gao, L.-Q., and Wydell, T. N. (2010). The visual magnocellular pathway in Chinese-speaking children with developmental dyslexia. *Neuropsychologia* 48, 3627–3633. doi: 10.1016/j.neuropsychologia.2010.08.015

Wang, X. L., and Tao, B. P. (1993). *Chinese Character Recognition Test Battery and Assessment Scale for Primary School Children*. Shanghai: Shanghai Education Press.

World Health Organization (2011). *International Statistical Classification of Diseases and Related Health Problems - 10th Revision*, 4th Edn. Geneva: World Health Organization.

Yang, Y., and Bi, H.-Y. (2011). Unilateral implicit motor learning deficit in developmental dyslexia. *Int. J. Psychol.* 46, 1–8. doi: 10.1080/00207594.2010.509800

Yang, Y., Bi, H.-Y., Long, Z.-Y., and Tao, S. (2013). Evidence for cerebellar dysfunction in Chinese children with developmental dyslexia: an fMRI study. *Int. J. Neurosci.* 123, 300–310. doi: 10.3109/00207454.2012.756484

Yang, Y.-H., Yang, Y., Chen, B.-G., Zhang, Y.-W., and Bi, H.-Y. (2016). Anomalous cerebellar anatomy in chinese children with dyslexia. *Front. Psychol.* 7:324. doi: 10.3389/fpsyg.2016.00324

Zhang, P., Wen, W., Sun, X., and He, S. (2016). Selective reduction of fMRI responses to transient achromatic stimuli in the magnocellular layers of the LGN and the superficial layer of the SC of early glaucoma patients. *Hum. Brain Map.* 37, 558–569. doi: 10.1002/hbm.23049

Zhang, P., Zhou, H., Wen, W., and He, S. (2015). Layer-specific response properties of the human lateral geniculate nucleus and superior colliculus. *NeuroImage* 111, 159–166. doi: 10.1016/j.neuroimage.2015.02.025

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership