

# FLUENCY AND READING COMPREHENSION IN TYPICAL READERS AND DYSLEXICS READERS

EDITED BY: Simone A. Capellini and Giseli D. Germano

PUBLISHED IN: Frontiers in Psychology and Frontiers in Education



# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-415-0

DOI 10.3389/978-2-88945-415-0

## 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](mailto:researchtopics@frontiersin.org)

# FLUENCY AND READING COMPREHENSION IN TYPICAL READERS AND DYSLEXICS READERS

Topic Editors:

**Simone A. Capellini**, São Paulo State University “Júlio de Mesquita Filho” (UNESP), Brazil

**Giseli D. Germano**, São Paulo State University “Júlio de Mesquita Filho” (UNESP), Brazil

Reading involves decoding and comprehension components and, to become efficient, it requires a large number of cognitive and linguistic processes. Among those, the phonological awareness, the alphabetic principle, the decoding, the fluency, the lexical development and the text comprehension development. The reading comprehension is strongly related with the development of vocabulary, oral language, linguistic skills, memory skills and ability to make inferences, and the world experiences of each individual. These processes become important only when the professional needs to deal with students presenting difficulties in learning how to read.

The difficulty using the knowledge of conversion rules between grapheme and phoneme to the word reading construction characterizes the dyslexia, which is a specific learning disorder with a neurological source. These difficulties presented by students with dyslexia interfere in their learning process impairing the learning development.

Knowing and following the reading development and its processes, as well as obtaining the punctuation of fluency abilities and students comprehension allow us to understand what happens when the student presents difficulties to read. This could help in the identification of learning disabilities and in the development of intervention programs.

**Citation:** Capellini, S. A., Germano, G. D., eds. (2018). Fluency and Reading Comprehension in Typical Readers and Dyslexics Readers. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-415-0

# Table of Contents

- 05    *Grasping the Interplay between the Verbal Cultural Diversity and Critical Thinking, and Their Consequences for African American Education***  
Horace Crogman
- 21    *Screening Protocol for Early Identification of Brazilian Children at Risk for Dyslexia***  
Giseli D. Germano, Alexandra B. P. de C. César and Simone A. Capellini
- 34    *Effects of a Syllable-Based Reading Intervention in Poor-Reading Fourth Graders***  
Bettina Müller, Tobias Richter, Panagiotis Karageorgos, Sabine Krawietz and Marco Ennemoser
- 43    *Sensitivity to Stroke Emerges in Kindergartners Reading Chinese Script***  
Su Li and Li Yin
- 51    *Applicability of the Compensatory Encoding Model in Foreign Language Reading: An Investigation with Chinese College English Language Learners***  
Feifei Han
- 64    *Contribution of Word Reading Speed to Reading Comprehension in Brazilian Children: Does Speed Matter to the Comprehension Model?***  
Alessandra G. Seabra, Natália M. Dias, Tatiana Mecca and Elizeu C. Macedo
- 71    *Reading Fluency as a Predictor of School Outcomes across Grades 4–9***  
Lucia Bigozzi, Christian Tarchi, Linda Vagnoli, Elena Valente and Giuliana Pinto
- 80    *Silent Reading Fluency and Comprehension in Bilingual Children***  
Beth A. O'Brien and Sebastian Wallot
- 92    *The Semantic Associative Ability in Preschoolers with Different Language Onset Time***  
Dina Di Giacomo, Jessica Ranieri, Eliana Donatucci, Nicoletta Caputi and Domenico Passafiume
- 100    *Reading Comprehension Assessment through Retelling: Performance Profiles of Children with Dyslexia and Language-Based Learning Disability***  
Adriana de S. B. Kida, Clara R. B. de Ávila and Simone A. Capellini
- 113    *Orthographic Reading Deficits in Dyslexic Japanese Children: Examining the Transposed-Letter Effect in the Color-Word Stroop Paradigm***  
Shino Ogawa, Masahiro Shibasaki, Tomoko Isomura and Nobuo Masataka
- 120    *On the Development of Parafoveal Preprocessing: Evidence from the Incremental Boundary Paradigm***  
Christina Marx, Florian Hutzler, Sarah Schuster and Stefan Hawelka
- 133    *Anomalous Cerebellar Anatomy in Chinese Children with Dyslexia***  
Ying-Hui Yang, Yang Yang, Bao-Guo Chen, Yi-Wei Zhang and Hong-Yan Bi

- 142** *The Role of Reading Fluency in Children's Text Comprehension*  
Marta Álvarez-Cañizo, Paz Suárez-Coalla and Fernando Cuetos
- 150** *Preliminary Validation of FastaReada as a Measure of Reading Fluency*  
Zena Elhassan, Sheila G. Crewther, Edith L. Bavin and David P. Crewther
- 160** *Discrete versus Multiple Word Displays: a Re-analysis of Studies Comparing Dyslexic and Typically Developing Children*  
Pierluigi Zoccolotti, Maria De Luca and Donatella Spinelli



# Grasping the Interplay between the Verbal Cultural Diversity and Critical Thinking, and Their Consequences for African American Education

Horace Crogman<sup>1,2\*</sup>

<sup>1</sup> Physics, College of the Desert, Palms Desert, CA, United States, <sup>2</sup> Research and Development, The Institute of Effective Thinking, Riverside, CA, United States

## OPEN ACCESS

### Edited by:

Michael S. Dempsey,  
Boston University, United States

### Reviewed by:

Alyse Jordan,  
Nova Southeastern University,  
United States

Gloria Victoria Jackson,  
Loma Linda University,  
United States

### \*Correspondence:

Horace Crogman  
hcrogman@gmail.com

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Education

**Received:** 02 April 2017

**Accepted:** 16 November 2017

**Published:** 11 December 2017

### Citation:

Crogman H (2017) Grasping the Interplay between the Verbal Cultural Diversity and Critical Thinking, and Their Consequences for African American Education. *Front. Educ.* 2:64. doi: 10.3389/feduc.2017.00064

The role language has in human learning has been discussed in the context of its impact on culture through African American communities. A strong link between thinking and language through the framework of question asking was reported. This essay improves upon Crogman and Trebeau's (Crogman and Trebeau Crogman, 2016) Generated Question Learning Model by incorporating language and comprehension as major tenets. The proposed argument is centered on language as the determinant of structured thinking, which in essence brings about learning through sensory experience. Further, the case was made for issues that may emerge in learning when the cultural norms of the learner are ignored.

**Keywords:** language, critical thinking, curiosity, learning, comprehension, culture diversity, dyslexia

## INTRODUCTION

Historically, the idea that language could shape thinking was considered erroneous and untestable (Bloom and Keil, 2001). However, with the input of a few decades of research in modern fields such as linguistic, sociology, psychology or anthropology, we have learned that people who speak different languages do indeed think differently and that even flukes of grammar can profoundly affect how we see the world (Levinson and Wilkins, 2006; Weiler, 2015). Thus language shapes our experiences of the world much like Piaget who considered children as entities able to build elaborated models of their environment by evolving from low to high-level conceptual prototypes (Inhelder, 1978). Vygotsky's views add to the discussion, by pin pointing to when the integration of speech and thought culminates, around 2 years of age, where infants become able to transfer language to their internal thinking, making their cognitive process more rational (Vygotsky, 1978). Thus, we contend with many that thinking and learning are inseparable systems; and thinking is facilitated through language, which is fundamentally the most important concept for human learning (Halliday, 1993; Weiler, 2015). Chomsky (1956, 1975) suggests that thought and speech are largely separated and argues that, in humans, thought is depending on specific cognitive domains. Chomsky believed in the innateness of the language ability. He proposes that our ability to use language, operationalized in some brain specialized regions (akin to module or domain specific perspectives) (Inhelder, 1978;

Cowie, 1997; Devitt, 2006), is based on our understanding of its mechanisms, which guides our speech and understanding of others' speech (Inhelder, 1978). Thus the implication is that language should have measurable effects on learning.

We can infer, from a number of critical thinking models, that they all suspect language to have a role in thinking, however, it is rarely clearly outlined. Indeed, despite a tremendous amount of research conducted on language or critical thinking, the interplay between the two has received less empirical attention (Romano, 2007). For example, Bloom et al. (1956) developed a model to classify the various levels of thinking complexity, but absent from this model is the role of language at each or either of these stages. Yet, language has been found to have particular effect on memory, perception, problem-solving, and judgment which in essence confirms the language-thought relationship (Hardin and Banaji, 1993). Zlatev and Blomberg (2015) may have successfully revived the language-thought relation hypothesis (postulating that thought and thinking take place in a mental language) by arguing: (1) for the disentanglement of language and thought; (2) that language from culture and social interaction can be unraveled; and (3) that all forms of linguistic influences are possible.

Finally, among the major theories explaining development through the lenses of experience and the human biological blueprints, Vygotsky (1978) and Rogoff and Lave (1984) posit yet in addition, that no development could happen without the sociocultural fabric we are born into (rearing, society's norms, language, traditions, etc.). Such views have been validated by the important cultural gaps observed between different ethnic groups within or across different continents. Thus, in the quest for best practices when it comes to conceptualize learning, the argument of this article also contends that the cultural fabric permeating through language must be taken into account and fundamentally understood, in any thinking/learning model, to be truly representative of the diversity of learners. Interestingly, culture does shape how we learned and it is also where implicit biases are formulated. Piaget (1976) note, "...everything suggests that, on discovering the values accepted in his immediate circle, the child felt bound to accept the circle's opinions of all other national groups." It is through the critical thinking processes that we must identify and overcome the impacts of stereotypes and biases. However, do culture-language dynamics take our critical thinking process hostage, especially in light of research that has shown its clear impacts on our judgment, analysis, and decision-making?

It is the goal of this essay to argue that language, a unique human communication system, is central to our experience, and appreciating its role in constructing our mental lives, brings us one step closer to understanding the very nature of human learning. We argue that we must fundamentally consider the fact that all theory of learning must have language as a tenet to be considered truly encompassing of the issue of learning. We will further in this essay explore connections between language and human curiosity, language as used in instruction, and the impact of culture within the dynamics of learning, language, and instruction. We seek to understand in what ways and to what degree language affects cognition, and how lacking or struggling with

language development creates more or less impeding deficits for learning and thinking.

## THE INTERPLAY OF LANGUAGE AND THINKING IN THE LITERATURE

Some scholars argue that specific word items in a given language influence how the mind splits reality into different categories, while others have proposed that the thoughts amalgamate into larger complexes through syntax (Bloom and Keil, 2001). The well-known Sapir-Whorf hypothesis (Whorf, 1956; Kay and Kempton, 1984) states that the structure of language determines and greatly influences the modes of thought and behavior characteristics of the cultures in which it is spoken. This hypothesis also suggests that certain thoughts in one language cannot be understood by speakers of another language.

For example, Russian speakers are better than English speakers at distinguishing colors, while Japanese speakers tend to group objects by material rather than shape unlike any other groups (Weiler, 2015). This shapes how people from different cultures orient themselves in space or influences how they process color. The Aboriginal community defines space relative to the observer, which means that a speaker would not be able to express themselves properly, or even get past a greeting if they are not constantly being oriented in space.

In this sense we ask, does language become a vehicle for the growth of new concepts, which were not in the mind, and perhaps could not have been there without the intercession of linguistic experience? The possibility that language is a central medium for concept formation has captured the interest of many linguists, and educators alike. Concepts are core foundations of thinking. They are grouping strategies to allow human beings navigate and understand their world as concepts are held in memory and help us in every day decision-making. Language then helps use create concepts for our communication.

A great body of evidence is suggesting that language influences conceptual development in humans (Markman and Hutchinson, 1984; Waxman and Kosowski, 1990; Boroditsky, 2001). This is illustrated in situations where individuals lack language, the progress of learning is impeded (Spelke and Hermer, 1996). It does seem that language affects our on-line perception of the world, shaping the categories we form, enabling us to perform logical inferences, and causal reasoning. For example, as Bloom and Keil (2001) argue, language brings about social reasoning, and structures the basic ontology about time, space, and matter. The character described in Schaller (1991), "Man Without words," did not have language for 27 years, and subsequently could not say anything about what it was like during this "wordless" time. A similar experience is reported in stroke victims' experiences after losing language to trauma (Sinanović et al., 2011).

In this article, we examine the interplay between language and thinking as discussed in Crogman and Trebeau Crogman (2016) learning model. This model establishes how to reach *critical thinking*, which involves thinking reflectively and productively, and evaluate evidence. Such thinkers have a tendency to be



creative, open to new information, and aware of more than one perspective. Language is very functional in expanding children's curiosity, reasoning ability, creativity, and independence (Conley, 2007b). Students who engage in critical thinking uses questions as tools to gain quality information that will help in making good judgments and decisions. Thus the question becomes how does language aids this process to make us better critical thinkers, or to say differently, efficient concept jugglers.

## THE ROLE OF LANGUAGE IN LEARNING AND CRITICAL THINKING ENGAGEMENT

How well we ask questions is based on how well the language in which we think is developed. If language does refine human thinking then we cannot escape the fact that it must play a pivotal role in learning theory. How does language make us better thinkers? Crogman et al. (2015) started by initially connecting thinking, question asking and learning by showing that: "Description invites students to ask 'what,' 'when,' 'who,' whereas analysis focuses on 'why' and 'how,' and evaluation encourages students to think beyond the phenomenon by going deeper and asking 'what if?'" The ability to question at increasingly complex levels refines the learning experience.

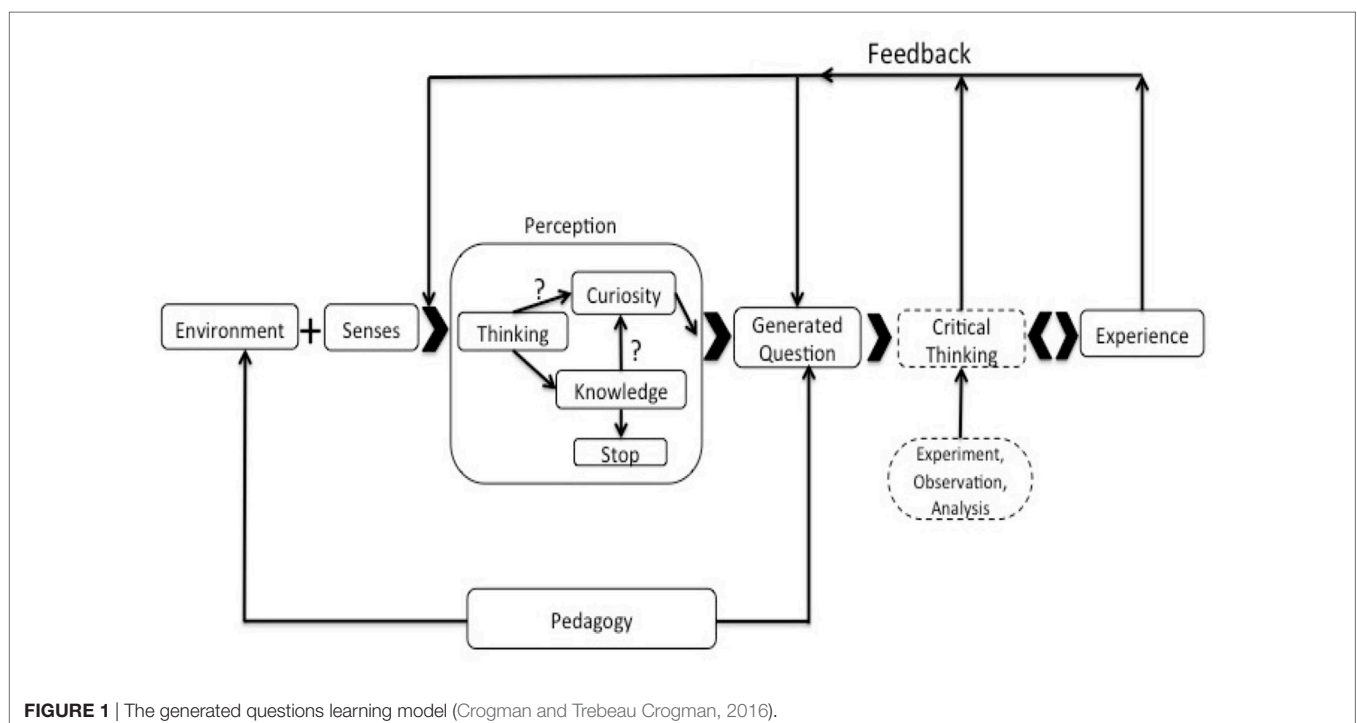
Crogman and Trebeau Crogman (2016) illustrate the interplay between students, the environment, and the educational practitioner (**Figure 1**). The practitioner uses pedagogies to influence the environment and awaken the student's curiosity, which in turn causes questions to arise in the mind of the learner. Questions naturally lead to inquiry, and inquiry leads to critical thinking, causing the learner to apply old knowledge or create

new ones. In this context Crogman and Trebeau Crogman (2016) suggest that learners that may have less access to expressive language, may internalize their questions creating challenges in the feedback loop that should be happening between thinking, questioning, and learning.

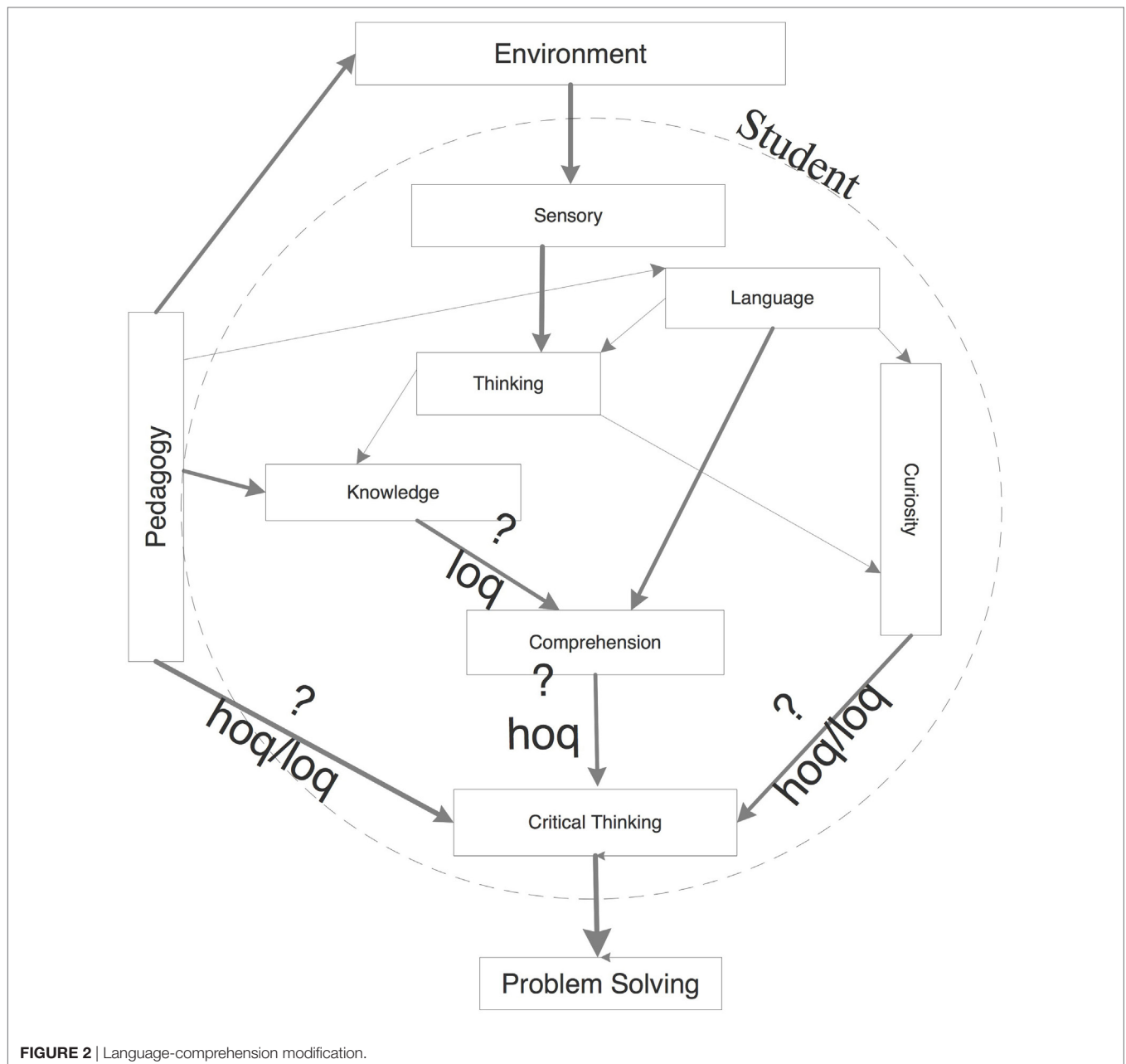
**Figure 2** makes two important additions to the learning model of **Figure 1**, these modifications are essential to critical thinking, and often overlooked: the addition of language and comprehension. For the most part, the importance of language on learning is known but the truth is that it is essential in the comprehension of knowledge as well. Language allows thinking (Tversky and Kahneman, 1985; Pelham et al., 2002; Boroditsky, 2003; Pica et al., 2004), and thinking allows question asking (Crogman and Trebeau Crogman, 2016), but the relationship between language and comprehension has largely not been discussed in learning models. How do they influence the learner's curiosity? In what way do they help in the critical thinking process?

To address these two questions we must see language as not a domain of human knowledge (except in the special context of linguistics, where it becomes an object of scientific study), but as the essential condition of knowing, the process by which experience becomes knowledge (Halliday, 1993). Further, as we are seeking to understand and to model how we learn, we should not isolate learning language from all other aspects of learning. Language in essence serves as the "signifier" for higher-level systems of meaning such as scientific theories (Lemke, 1990; Martin, 1991) and is a prototypical resource for making meaning (Halliday, 1993).

When there are difficulties in the process of language development, there may emerge neurobiological problems





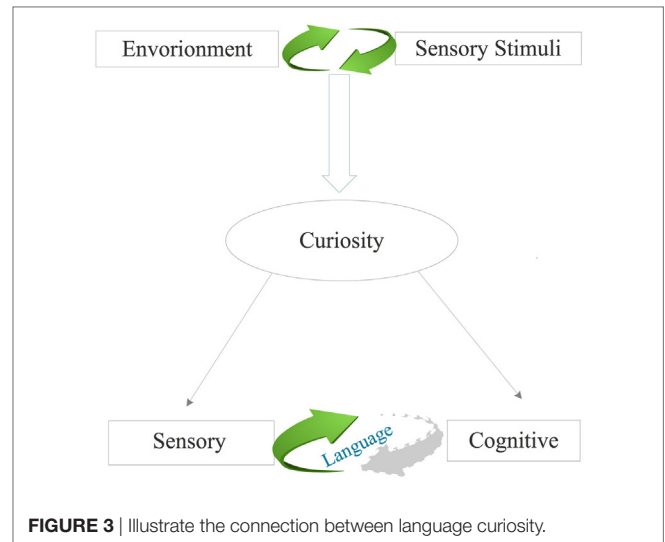


such as dyslexia or reading and comprehension deficiencies. Comprehension involves building meaning from language (Sparks, 2012). The ability to make meaningful connections across contexts helps textual and discourse comprehension. However, prior to that skill comes the need for having developed basic knowledge about those contexts and other more general facts. That base allows developing memories, which in turn informs on the contexts illustrated. Comprehension also requires a fairly automatized phonological process of binding and separating components of language, into units that can constitute knowledge to be stored (Sparks, 2012). Language, serves as a set of processing cues or instructions that guide construction of memory for discourse (Gernsbacher, 1990; Givón, 1992).

The nature of language is also combinatory (Spelke, 2010), thus the learner's language ability consists of a basic level (e.g., decoding and fluency) and higher order processes (e.g., the ability to make inferences). This requires the learner to possess a rich vocabulary, oral language skills, and reading skill (Sparks, 2012). The better the learner's language development, the more successful is their comprehension. This means that success in comprehending larger units of knowledge requires that learners make inferences to connect ideas both within and across local and global discourse contexts. Sparks (2012) explains that, "prior knowledge is crucial for disambiguating concepts, making predictions, and inferring unstated connections among ideas." Thus, comprehension is directly related to thinking

(Aloqaili, 2012) because it pushes the reader to reflect on prior knowledge to apply it or create new ones. Therefore, successful comprehension will result in the learner retrieving, updating, manipulating, and applying knowledge in order to ask questions and solve problems. Young infants, whom language is not yet well developed, rely more heavily on thoughts and action with impulsivity rather than rationalization. Gradual sophistication of our language ability helps us to think: that is, to logically reason about the world, while we also develop *inhibitory control*: “the ability to ignore distractions and stay focused, and to resist making one response and instead make another” (Diamond, 2006). Increase in inhibitory control aids children in regulating their emotions, and, behavior and helps them become more effective problem solvers. Therefore, good thinking is related to how well the learners comprehend, and allows them to construct meaning. Getting a sense of language is not based solely on syntax or word meaning understanding, but on understanding what is intended when those words are put together. Children with comprehension deficit experience show weakness in processing written and oral language, higher order thinking skills, and visual and auditory memory. Crogman and Trebeau Crogman (2016) argued that such deficit can be corrected through question asking, since it is so basic to understanding and learning. A pedagogy that involves learners in the skill of asking questions will improve their comprehension, which is directly associated with language development. Through that feedback process (Figure 1) question asking directly impacts comprehension and language.

Sparks (2012) points to the fact that prior knowledge, which includes information recently activated in short term memory (e.g., previously mentioned text concepts), as well as the personal experiences, facts, ideas, and understandings stored in long-term memory, is the most critical variable. As such, in Figure 1, a change in the learner’s environment is sensed through their sensory receptors, which brings about a response. Crogman et al. (2015) deconstructed the connection between environment and thinking processes showing how changes in the environment evoke perception and provoke responses, which are the result of thinking. They point out to two possible outcomes: (1) thinking creates a question that is answered by memory content. This is illustrated in Figure 1 by the arrow that goes to knowledge. Thus, from prior knowledge, curiosity is awakened by the question generated (question mark), otherwise, the path ends caused by lack of interest (Stop). It is comprehension of this prior knowledge that determines the direction of learning. Kandeou et al. (2003) highlighted the importance of considering the influence of anterior knowledge in the construction of relations between concepts and ability to comprehend and predict language. These are examples of skills to acquire, at different levels of language development complexity (e.g., from simple decoding to inferring) to be able to communicate effectively both orally and in writing or reading. In Figure 2, comprehension of prior knowledge requires language development of the learner. This analysis is normally on *low order questions* asking. (2) when thinking does not meet prior knowledge, curiosity is aroused, and pushes the generation of new questions (see also Loewenstein, 1994). Curiosity drives both low and higher order question (hoq)-asking mechanisms. Language clearly impacts cognitive curiosity as shown in Figure 3.



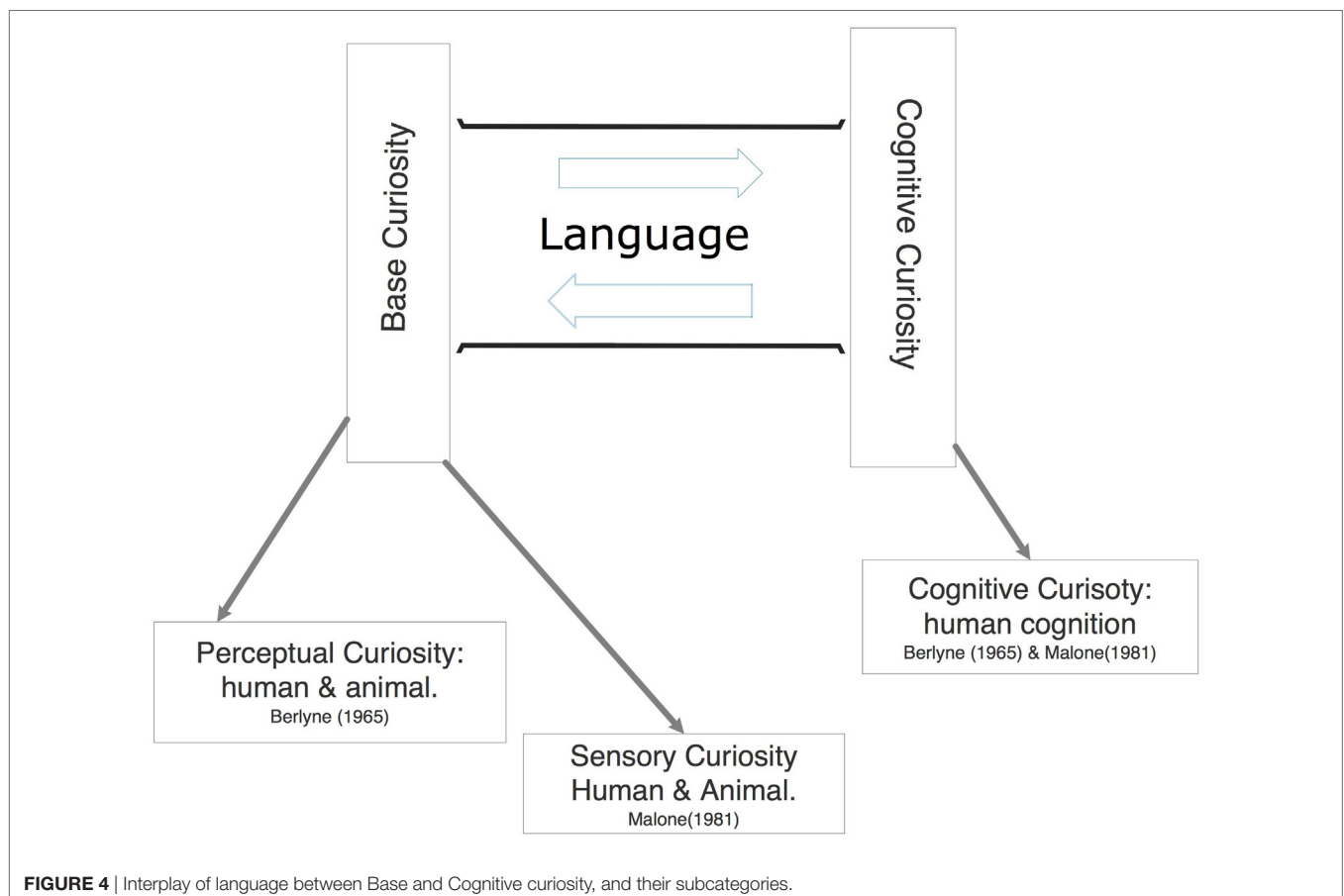
The learning model illustrated in Figure 1 requires the educational practitioner to use pedagogies to entice the learner’s engagement. That is, drawing the learners’ curiosity in. Curiosity then becomes the first stage to inquiry and/or critical thinking. Both Berlyne (1954, 1966) and Malone (1981) divide curiosity in two stages: *perceptual* or *sensory* which is present in all animals and humans, and *cognitive* which is a human-only domain. We address here the implications of this difference in the interaction of language and learning. In Berlyne’s (Berlyne, 1965) model, perceptual curiosity arises from conceptual conflict, which then morphs into epistemic (related to knowledge) curiosity, through question asking. Malone’s model starts with the sensory curiosity, which is aroused through environment as in Figure 1, to bring about cognitive state processes. Loewenstein (1994) suggests that curiosity is the intersection between cognition and motivation, which manifests cognitive induced deprivation as result of a perception gap in knowledge and understanding. Loewenstein did posit the idea of the “information-gap” perspective, which states that, in order for curiosity to be present, the learner must already have some level of knowledge. Chomsky’s (Chomsky, 1956) model suggests this as well in his exploration of the existence of some form of language in infants. As we explained, since language helps to formulate concepts, an, as Loewenstein (1994) suggests, infants’ curiosity is aroused by cognitive conflict, then such conceptual conflicts are factor that could facilitate student learning. This stems from the incompatibilities between symbolic responses and the conflict engendered by them (Berlyne, 1960). Berlyne thought that it must underlie the notions of truth and falsity, which can only be achieved if there is prior knowledge as Chomsky infers.

We push back a bit here to suggest that curiosity, at its basic stage, is found in all animals and is purely sensory. Meaning that it is not based on any prior knowledge. For example, a newborn may be curious about a colorful stimulus without having prior knowledge or concept about color to begin with. However, cognitive curiosity is based on prior knowledge or concept, and arises from conflict in information or unresolved

stimuli interaction. Cheney and Seyfarth (1998) speculate that animals lack language for the following reasons: no rudimentary theory of mind, and no ability to generate new words, and syntax, which are all present in young children. Animals have a number of in-born qualities they use to signal what they feel, but these are not like the formed words we see in the human language. Thus it is reasonable to conclude that animals could only have a notion of concepts if their communicative gestures were primitive forms of language; this being said, we must realize that Berlyne's perceptual curiosity is developed after sensory interaction with the environment which births the communicative gestures seen in both animal and children. Therefore, Malone's conception of curiosity must be the first stage before knowledge is acquired. This difference in cognition between humans and animals is experimentally verified. Kalia et al. (2008) point out that, in both animals and humans, there is categorization of non-geometrical modules (concrete concept or object such as a rock; allowing one to compute orientation in relation to a wall), and geometrical ones (abstract concept or object such as color concrete concept or object such as a rock; allowing one to compute orientation in relation to a wall). In animals as well as in newborn to toddlers these modules do not speak to each other. Fernyhough (2008) argues that humans are able to integrate geometric information (the short wall on the right) with the non-geometric information (the blue short wall,

not the white). It is believed that this is the result of language development in humans. Therefore, learning is driven by question asking, which leads to further inquiry behaviors. How does question asking play such an important role?

Because language does impact and streamline thought, it must affect a child's curiosity, leading to good question asking behavior. There is a clear development in the learner's ability to formulate questions in response to their curiosity development (Crogman and Trebeau Crogman, 2016). To formulate questions language is important. We speculate that there is not a direct correlation between sensory curiosity (i.e., "*curiosity base*"—Figure 4) and language. Animals exhibit curiosity even though they do not have language. Spelke and Hermer (1996) speculate that one of the main differences between humans and animals is the human formulation of language. They compared children (newborn to toddlers) and rats, on diverse tasks and their findings indicate that children deviate significantly from rats at about age six (Hermer-Vazquez et al., 2001), a point at which they are able to express complex language within their now fully integrated cultural norms. In human adults reorientation exercises are solve easily that is they quickly find an object left of a blue wall (Hermer and Spelke, 1996). Spelke (2010) propose that this ability emerges in synchrony with the development of spatial language such as expression of left or right terms, and is well known in developmental studies.



Another case study described by Schaller (1991), when the subject did not have language, he still exhibited curiosity, even though his ability to think was somewhat impeded. The transition to language in toddlers is what correlates to Berlyne's (Berlyne, 1954, 1965, 1966) and Malone's (Malone, 1980) second aspect of curiosity. From toddler to preschool where the learner's access to language is facilitated, more basic questions can be asked to aid their exploration (Borowske, 2005). Language develops the cognitive process in humans, how exactly this is done is up for debate. The link between cognition and language was proposed by Chomsky (1956), who believed that children are born with specific language acquisition devices and linguistic knowledge. The more accepted view today is centered on learning and not on innate structures (Harris, 2006). Piaget emphasized the commonalities between language and cognition proposing that language emerged out of the same broad cognitive changes that transform the sensorimotor processing of infants into formal and logical mind of adults (see **Figure 3**). Cognitive research has led to the idea that both language and cognition have complex similarities and differences influenced by genetic (Chomsky, 1956), environmental input (Elman et al., 1996), and cultural learning factors (Harris, 2006). It is through these basic questions that the critical thought process is engaged. Once comprehension has occurred (**Figure 2**), the learner generates *hoq* asking.

**Figure 4** proposed to divide curiosity into Cognitive and Base Curiosity where Base is further, divided into two parts: *sensory* and *perceptual*. We propose that perceptual curiosity results from the effect of the cognitive on the sensory. Thus, because of language, there is a constant interaction between the sensory and the cognition as shown in **Figure 3**. The back arrow (**Figures 3** and **4**) is fainter to represent that cognition's influences base curiosity through language. We can reason that because language is interconnective between the geometric and the non-geometric modules, it gives rise to specific cognitive processes in humans (represented in forward arrows in **Figures 3** and **4**). Cognition then is a reflective interplay in thought process to thinking in a coherent and sequential manner.

Further, seeing the result when humans seem not to have language or have a deficit in it, this suggests that language influences thinking in very profound ways. Language may not completely determine thoughts but it is clear that it streamlines thoughts and strongly influences thinking. Recent work from Zlatev and Blomberg (2015) supports that idea in their disentanglement of language from thought and culture to show that it is fundamental to human learning. For the most part the human thought process seems very random; for example a young child may experience an electrical shock by sticking an object into a plug or get burned by the stove. The child learns from this terrible experience, which prevents future repeats as this is stored in memory. This response may be completely behavioral to begin with, but here the pathway commences where the child uses language to formulate conceptual understanding in the mind and this development continues such that the child's curiosity transitions beyond mere behavioral. The child's experience of learning lends to the blank check theory of learning by Locke (1975) and Piaget (1976) that children learn gradually from their environment, in which every experience builds a set of *a priori* knowledge for the next

(Crogman and Trebeau Crogman, 2016). Communication at this stage is limited, making it more difficult to communicate what has happened with clarity or express how they felt; it is here that language helps the human thought process in order to convey feelings and perceptions of the world around.

How to make sense of the following experiment performed by Hermer-Vasquez et al. (1999)? Adults participated in a reorientation task where they listened to a tape recorded prose passage and repeated it continuously, word for word (verbal shadowing); they were observed to lose the ability to combine geometric and non-geometric information and performed like rats and children tested on the same task. Their thought process was "foiled," and word sense disconnected by the temporary lost of language. When the experiment was repeated with a second group using a different task by listening to a tape-recorded percussion sequence and repeating the sequence by clapping (rhythm shadowing), the adults were able to combine geometric and non-geometric information. The researchers concluded that natural language helps in the construction of new spatial concepts and their active use. We see this as evidence that language is strongly correlated with thought and determines structured thinking (the ability to organize thought logically).

Furthermore, one finding of cognitive research is that curiosity tends to decrease with age because children become cautious (Hutt and Bhavnani, 1972), however, language development continue over the life of the learner. Since the result of good language development makes the learner a better questioner, then question asking stands as a method to counter the effect of declining curiosity (Crogman and Trebeau Crogman, 2016). Questioning can also be used to deepen and enrich knowledge as well as expand the understanding of content. Question asking helps learners link all prior knowledge, think about the exact content, draw out meaning in order to make coherent explanations, develop inference skills, and construct key points to build mental representations (Martin and Duke, 2011; Crogman et al., 2015). Therefore, language is directly linked to question asking. Further, we use language to think aloud, which is an effective comprehension strategy that requires the learner to extract, construct and think about the content, which facilitates knowledge. It taps into a metacognitive process where learners monitor their reading before, during, and after reading (Baker 2009). The end goal for the learner is to become better at critical thinking to effectively solve problems, this skill is inevitably based on how developed language is.

Thus, we outlined a prior learning model, which has taken into account important aspects of learning, which are the environment and how educators manipulate it, their ability to arouse curiosity, and to drive learning by teaching how to ask questions. In that first model however, language and comprehension have been overlooked. In **Figures 2–4**, are illustrated the role of language in the skill of learning. We see the importance of developing language skills to operate such critical skills as thinking and asking questions, without which learning is not possible. These skills separate us from all animals, and must be taken into account in any Learning Model, using language. The issue is, how can educators operate such strategies to expand learners' horizon when there are language barriers in their students (outside of ethnic foreign language barriers)? Indeed what



does happen when language is misunderstood, and what is the impact of such a problem in the learning processes of developing learners?

## LANGUAGE AND CULTURE'S INFLUENCES ON CRITICAL THINKING

Human cultures provide the framework in which languages develop, and influences how they are used and interpreted. In some groups more than others, gestures, glances, changes in tones, along with other devices are widely used to emphasize what is communicated. Language is closely related to culture, but in reality its influence is often overlooked (Hadley, 2000). Nida (1998) suggests that language and culture cannot exist without each other, and languages not only represent elements of culture, but also serve to model culture. If the influence of culture on language is ignored however, serious misunderstandings will emerge in communication. Nida (1998) proposes that words are determined by both syntagmatic and cultural contexts, but language still may change in word meaning faster than it changes culture itself.

Ricci and Huang (2013) argue that cultural influences do affect thinking styles, shape personal thinking preferences, and have their grip on critical thinking strategies since it has been shown to affect individuals' thought processes, judgment, and decision-making and inhibit the ability to be unbiased. There are no empirical data found in the literature that addresses the issue of the interplay between culture and critical thinking directly. Yet, in the context of learning and thinking contexts, such as education for example, the place of culture and language is important as culture comes with bias in thinking. In that pursuit, a number of researchers (Paul and Adamson, 1990; Ennis, 1998; Ricci and Huang, 2013) argue that the ability to address bias is an important dimension of critical thinking. How can this be true if thinking in itself has fallen prey to cultural conditioning? If language and culture impact an individual's thinking, does it mean that critical thinking, which is a tool for overcoming biases is inherent with them?

Language is the vehicle through which we often experience cultural biases. This is a preference or an inclination that inhibits impartiality; prejudice (American Heritage Dictionary, 1983), or "a predisposition or a preconceived opinion that prevents a person from impartially evaluating facts that have been presented for determination. A bias held limits the critical thinking processes" (West's Encyclopedia of American Law, 2005). The issue with most biases is that they can become unconsciously activated. To illustrate, a new White teacher, working in the predominantly Blacks and Hispanics South Bronx said one day in class, "...And for homework, I'm going to give you people..." The reaction of the students immediately turned to anger. The incident became a teaching moment when a student asked, "What do you mean by You People? We don't like to be called You People!" to which the teacher apologized.<sup>1</sup> Such is a perfect example of cultural

misunderstanding attached to language. As we will detail further, good critical thinking skills helps us to examine our biases. It is here that the language-culture dynamic exerts its influence on thinking, which could be very harmful to the learning process, meaning that critical thinking in itself is subjected to cultural influences, which causes thinking in itself to be shaped into such biases.

Levinson and Majid (2011) by looking at the differences in the thinking processes associated with the type of language spoken found that language and culture influence cognition. Such data are evidence that language is used to form concepts and categories, which are born by culture, and influenced by their specific rules and choices in language usage. For example, Davidson's (Davidson, 1994) argues that in Japanese culture, critical thinking is inhibited due to a number of cultural demands, which do not encourage diversity of opinions as most of their education processes are based on rote memorization.

In the United States, culture strongly influences the education system through which policy and instruction are formulated. An impressive body of research spanning decades addresses such issues and leads to some uncomfortable conclusions, yet, it is difficult to isolate effects of race and culture from other factors. The things we experience and observe in our culture or about other cultures compel us to create biased (meaning often unilateral or containing only partial information) concepts, categories, and stereotypes. Hamilton and Troler (1986) define stereotypes as positive and/or negative belief, expectations, and knowledge established about designated or singled out groups. Bigler and Liben (2007) along with a large number of researchers from diverse disciplines, posit that such categorizing is an innate cognitive behavior. They explain that humans gravitate toward this type of cognitive strategies by mere need to conserve mental energy, understand and predict the world, and reinforce the feeling of belonging that "ingrouping" and "outgrouping" affords. Some researchers explain that such biased "grouping" become so ubiquitous that the stereotypes attached to them are integrated into unconscious layers of our cognition. If it is unconscious, then those biases and stereotypes will be applied broadly without the use of more analytical thinking. The goal of the argument further is to examine the impact that negative stereotypes can have on the critical thinking process. Their influence on critical thinking is best reflected in the performance of Blacks and other minority groups. The focus will be the effect of stereotypes on learning and thinking in the African American learners.

Stereotypes, are cognitive constructions which encompass a set of convictions and assumptions that are presumed, in the case of "racial stereotypes" here, to be shared among members of a same racial group, often in a negative context (Jewell, 1993; Peffley et al., 1997). These stereotypes are played out in society at large in their influence on public policies and opinions against particular groups. The culture of the dominant Caucasian group in the US (we will call it here "the ruling class") has the largest representation in political and legislative decisions where these negative biases often appear in the formation of policy. Jewell (1993) argues that there is an obvious trend in American culture to discriminate against, and deny access to social institutions to Blacks. These stereotypes formulated in the language of the ruling class will continue their effects on such groups for generations. It

<sup>1</sup>Community Coalition On Race (n.d.). *Stories from Our Community about Language, Stereotypes, and Communication*. Available at: <http://www.twotowns.org/language, stereotypes, & communication.html>

is well documented that the ruling class tends to think negatively of Blacks: males are deemed violent and brutish, while females are seen as dominant, and lazy (Peffley et al., 1997). Blacks are considered to be inferior to all other groups, for example, they were believed to be mentally inferior physically, culturally unevolved, and apelike in appearance for centuries. Such absurd perspectives were well engrained in our historical highest institutions, infrastructures, and resources like the Encyclopedia Britannica published in 1884, stating authoritatively that “the African race occupied the lowest position of the evolutionary scale, thus affording the best material for the comparative study of the highest anthropoids and the human species” (Plous and Williams, 1995, p. 795). Contrary to common opinions, such views still exist today and are propagated in the educational system of the American classrooms in more or less subtle ways.

Recent research has shown that members of the ruling class are likely to hold these stereotypes especially with respect to issues of crime and welfare (Green, 1999). Welch (2007) points out that the ubiquitous stereotype of Black men as “criminal predators” is so engrained in society’s perception that it permeates the global unconscious to the point of affecting systems such as Justice or Law Enforcement, and influencing their practices and justifications for bias. Another example is the overrepresentation of Blacks as sports figures (Peffley et al., 1997). Edwards (1973) observed that the arguments from social Darwinism that helped solidify the stereotypes in American communities such as natural ability of Blacks and intelligence of Whites are still used as mutually exclusive attributes to account for racial differences in sports performance. An experiment by Stone et al. (1999) showed the impact of stereotypes on athletes’ performances. Black participants performed significantly worse than did control participants when performance on a golf task was framed as diagnostic of “sports intelligence”; on the other hand, White participants performed worse than did controls when the golf task was framed as diagnostic of “natural athletic ability.” What does this tell us about the influence of language on concept formation such as cultural stereotypes? What is the impact on classroom learning or development of critical thinking skills in some groups if the perception of the teacher and students are shaped by these stereotypes? The Clark Doll experiment (Clark, and Clark, 1939) illustrates the pervasiveness of racial bias and how early it seems to be engrained in the mind of children, who then grow up unknowingly categorizing according to these implicit biases. Children aged 6–9 were asked to choose a doll to play with, and also to indicate which one looked like them. It was found that black children often chose to play with the white dolls more than the Black ones. In another experiment (Davis, 2009), children were asked to tell which doll was the mean one and which doll was the nice one. Children overwhelmingly chose the Black doll as the mean one. The devaluing of the Black doll is evidence of racial biases as induced by many factors in society, and will impact individuals at every level of the social fabric. Since it is stereotypical belief that Whites are smarter than Blacks, this will affect how teachers perceive Black students, how Black learners perceive themselves, and how their peers perceive them. Correcting these stereotypes implies reformulation and acceptance of cultural language diversity, to restructure the belief

system attached to racist stereotypes that creates a false narrative in our youths.

The problem is twofold: instructors having a grasp of what cultural diversity really is and avoiding the pitfalls of stereotyping, and also understanding that that diversity comes with language hallmarks that may not sound or look like what they are accustomed to. The degree with which one tends to stereotype has been connected to the degree to which one holds the belief that people’s characteristics cannot change or tend to be the same and constant among certain groups (Levy and Dweck, 1999). McKown and Weinstein (2003), for example, show that these beliefs, crystallized into stereotypes can translate into behaviors that may impact children school performance. However, they argue that, with proper guidance and intervention, such tendencies can be reversed, and the deleterious effects of biases transformed. As argued by great thinkers of the beginning of the century such as DuBois (1903), in the case of African Americans, such stereotypes as born by the ruling class, have caused them to choose to close the door to higher education to these groups, thereby also stifling their opportunity to take part into the academic exercise of critical thinking. Countless research has established to date that such stereotypes about race have permeated education in harmful waves, and pushed individuals, who vowed to educate the masses, to close off opportunities of education to underrepresented groups even by their grading attitudes, expectations, and behaviors in class. Researchers analyzed educational, demographic, and survey data of 10,000 high school sophomores and their teachers using the Education Longitudinal Study of 2002, to show that teachers typically underestimating their students’ abilities, actually created a negative impact on their academic expectations of themselves, and this was especially harmful among Black students (Cherng and Halpin, 2016; Cherng, 2017). Further, Fleming (1984), and Smedley et al. (1993) along with a large body of recent research (Locks et al., 2008; Hurtado et al., 2009) demonstrated that racial biases encountered in school severely negatively impacted Black students’ academics, critical thinking, sense of belonging, and emotional development through heightened stress levels. However, they stress that the distress experienced by racism in school has a different impact on these students and creates unique sets of cognitive states unlike other regular sources of strains, pressures, and difficulties. Indeed, such concepts as “stereotype threat” (the belief about racial inferiority), have been coined to show how much of an insidious impact cultural stereotypes have created on the minds of those who are the aware victims of these issues, and how much of an effect these views have had on their ability, or even their beliefs about their ability to think and reason. The resulting impact is a negative effect on the development of students’ critical thinking due to teachers’ biased perception of students, and the students of themselves.

Diversity in the classroom is notoriously misunderstood, and a known source of miscommunication between educators and students. Thus, when it comes to the language differences brought by cultural diversity, a representative from an Asian cultural background may not experience society and educational system pressures through the same negative stereotypes as do Blacks, even through their English speaking may be strongly influenced by their culture. In another experiment, researchers attempted



to distinguish if distinction between stimuli such as colors was based on language or some other visual mechanism (Kay and Kempton, 1984). In their first experiment, they asked English speaking and Tarahumara speaking participants to explain which of three colored chips was the odd one in the context of their color distance, knowing that English speakers and Tarahumara speakers do not see color distance the same way. Expectedly both groups did not give the same distinctions between the three chips. To assess whether language was the reason for that difference, in another experiment the researchers eliminated part of the choice and also constrained the choice to how much *blueness* and *greenness* difference there was, thereby eliminating the color categorization afforded by language specificity. Surprisingly, English speakers aligned with the Tarahumara. Language was somehow a barrier to the two groups seeing eye to eye on color categorization. This experiment could potentially be expanded to other domains to highlight how language constrains perception, concepts, categorization and other vital skills necessary to communicate. Researchers showed that linguistic differences influence how speakers of two different languages view events. In one experiment German and English speakers were compared on ambiguous and goal-oriented scenes matching. German speakers matched twice as much as English speakers situations showing that they, more than English speakers, were focused more on people's actions outcomes than on the actions *per se*.

Thus the specificities of language are fundamental to many aspects of communication, and multilinguality adds another layer of complexity to the problem. Multilingual individuals are more advantaged in the classroom because speaking in other languages aids cognition and thinking processes at different levels (Kubota, 2013). Indeed, multilingual individuals have been found to present clear cognitive advantages and to be more easily flexible in their thinking. A large meta-analysis looking at over 6,000 bilingual individuals showed superior abilities for example in attention, memory, metalinguistic awareness, and understanding of symbolism (Adesope et al., 2010).

Multilingualism by nature also often affords multiculturalism, which can allow individuals to have more accepting attitudes toward others (Kubota, 2013). This begs the question why do Blacks tend not to be seen more positively and perform better in the classroom given their inherent multiculturalism? Could it be that multilingualism also constitutes a setback? For example, it may emphasize commonality and natural equality across racial, cultural, and gender differences for everyone, which then may perpetuate certain stereotypes such as Asian students being passive and silent learners, who fail to become autonomous learners (Zhao, 2008). In the case of Blacks, the uniqueness of the African American English language, strongly influenced by the African slave ancestry, and creole cultures for example (Green, 2002), bares the marks of a history that still holds prejudice on its shoulders without the respect due to its legacy. Thus in the classroom, as pointed out by Kubota (2004), ethnic customs and traditions are merely displayed and consumed without learning about their sociopolitical origin. Differences are ignored in this multicultural environment, which obscures the ruling class' power and privilege. Kubota (2004) argues that failing to appreciate this diversity of culture can only promote the continuation

of "racial and linguistic hierarchies." The impact of school and language should have given Blacks a much better foothold in the American society, but instead the whole culture still seems to be condemned and deemed as negative.

How do we mitigate this negative aspect of culture and language perceptions on individuals' success in learning? Kubota (2004) argues for "*critical multiculturalism*." By this she means that one must understand and appreciate the invention and performance of identity in intercultural communication, through examination of how groups construct their identity in social and historical ways (Kubota, 2004). This types of multiculturalism demands that both students and teachers, "critically examine how curricula, materials, daily instructions and social differences are constructed, legitimated and contested within unequal relations of power" (Zhao, 2008), a critical reflection about the discourses' power/knowledge and social impact.

The point here is that language and culture are so intertwined that when culture bares the burden of prejudice, so does language, and all learning attached to this context is highly impacted. Thus a better understanding of culture also implies a better understanding of language, and thereby improves learning and teaching directly and indirectly. Further, it requires a certain degree of ones' own awareness of one's biases in order to even start applying the critical thinking process to those biases, to remove them from our cognitive procedures when we apply our judgment to understand or know others. Thus, the data examined here show that under the pressure of certain stereotypical stressors, students' performances is negatively impacted, which is a reflection that their critical thinking processes are also affected.

## THE CASE FOR LANGUAGE AND DYSLEXIA: HOW BLACK CHILDREN CAN BE IMPACTED

As a Black, man coming from a cultural environment with a unique phonological fabric, I have had to face struggles between my understanding of language and my spelling of that language in an academic environment. Though I did not lack the understanding of the general English language, my own experience influenced my writing, and more often than not, this information eluding my professors (knowingly or unknowingly), caused me to receive grades that were at odd with my general understanding and cognitive abilities. Hence, a diagnosis of dyslexia seemed quite unfitted.

In reflection, cultural norms in many cases may be responsible for misdiagnosis of dyslexia in multicultural children. Too much of a broad brush is used to characterize this disability. Why is it necessary to consider such a question in this essay? From my point of view, a large number of reading disabilities may be a result of the interactions between language and culture. Especially in the context of the American culture's implicit biases, which are responsible for misdiagnosis and under diagnosis in Black populations (Robinson, 2013). A number of studies highlighted the unique difficulty associated with being Black, male, and dyslexic altogether (Robinson, 2013). Indeed, issues for this

population compound into a cluster of roadblocks associated with unfair treatment and lack of access to resources, while at the same time suffering from symptoms inherent to the reading disability itself which often has translated also into misplacement or placement into inappropriate special education support while the need is elsewhere (Catts et al., 2005; de Valenzuela et al., 2006; Vellutino and Fletcher, 2005; West-Olatuji et al., 2006; Gardner and Hsin, 2008).

Lyon et al. (2003) provide a definition of dyslexia, which is widely accepted and seems to capture the essence of this reading disorder:

Specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction.

Although language is impacted and impacts brain development, we must push back on this definition, which suggests that dyslexia is solely neurobiological in origin. The term “*neurobiological*” implies a sole biological origin for the disorder, which inherently ignores the impact of how culture shapes language, and the biases that may result in reading disabilities that could be acquired from the environment. Phonological deficits can result from the effects of the environment a child grew up in and may create reading difficulties as result. Moreover, dyslexia literally means—difficulty with words (Catts and Kamhi, 2005). Hudson et al. (2007) explained that, “People with dyslexia often have trouble comprehending what they read because of the great difficulty they experience in accessing the printed words.” The very fact that phonological issues define dyslexia, we can speculate that cultural norms and implicit biases that are a direct influence on language can cause a misinformed or unacquainted professional to mis- or underdiagnose. One of the first indicators that the question Dyslexia could be controversial in the diagnostic domain and how it contributes to achievement, is the relative lack of ethnically diverse dyslexia research in educational literature, and the failure to highlight the specific conditions of certain groups such as Black males with this developmental reading disorder. Biases may also inhibit professionals from developing an understanding of the resources and interventions needed to enhance the learning and academic achievement of these groups (Robinson, 2013). For example, the literature reveals that in general, research articles do not report conclusions by race, and also that there is a strong need for more reading interventions to include for example Black students (Lindo, 2006; Hoyles and Hoyles, 2010; Proctor et al., 2012). In this context, if dyslexia is not accurately diagnosed, Black males with dyslexia will continue to experience academic problems, be seen as defiant, and receive erroneous labels of emotional or behavioral disorder (Gardner and Hsin, 2008).

Language is fundamental to human learning, yet cultural stereotypes attached to language can counterbalance the effectiveness of learning as shown above. For example, teachers who exhibit explicit or implicit racial prejudice make recommendations to place Black males in dead-end situations that can lead

to frustration and alienation (Ford, 2010, 2013). Whiting (2009), shows that such views can influence students’ behavior, perhaps causing withdrawal from school, acting out, low self-efficacy, poor attitudes, and eventually low academic success among which low reading abilities are directly related to this cycle. Oftentimes words used in African American and Black cultures are often deemed inappropriate within educational contexts, and sometimes lead instructor to see students as having behavior issues, and poor language skills. There are phonological cultural specifics in the way Blacks pronounce some sounds differently from the White culture’s words use. This causes this group to get bad reading grades. Even gestures from the Black culture in the school or workplace may create or influence negative stereotypes. Since most Blacks in America are English speakers, some pronunciation of words are heavily influenced by the cultural background which can cause the classroom, outside of this culture, to perceive these students as ignorant. For example, “*cub*” sounds similar to “*cup*,” “*street*” is pronounced “*skrit*,” “*thin*” as “*tin*,” “*the*” as “*de*” or “*da*,” “*ask*” as “*aks*,” etc. (Green, 2002). Further, in the word “*sing*” *n* and *g* are combined (Green, 2002). The manner in which tenses are used, and grammars constructed produces language patterns that are very different from what is being taught in the American classroom, and as result, schools, like a foreign country, can become hostile and difficult environments to evolve and learn in. Understandably then, these unique pronunciations of letters and phoneme combinations can cause these students to spell some words differently.

Additionally due to marginalization from systemic racism, Black parents may often leave their children to fend for themselves in learning the language of the ruling class, being themselves not equipped to help their children in the same way as their White counterparts. Snow et al. (1998) argue that dyslexia is not caused by poverty, developmental delay, speech or hearing impairments, or learning a second language, but they admit that these conditions may put a child more at risk for developing a reading disability. However, considering the fact that language helps in the development of the brain, and that earlier years are important in child learning due to the brain’s plasticity, a delay in development can bring about a reading disability. Poverty is the medium through which reading disability could flourish. Marginalized communities tend to be poverty stricken, and children suffer from a number of learning disabilities. It may very well not be the cause *per se*, but could be part of the environmental circumstance, which should not be ignored.

Another major problem encountered is the norm-referenced, standardized testing tools by which achievement abilities are measured across the board, and blind to cultural dynamics. “African Americans currently score lower than European Americans on vocabulary, reading, and mathematics tests, as well as on tests that claim to measure scholastic aptitude and intelligence (Jencks and Phillips, 1998).” These tests alone do not accurately measure a student’s intellectual and academic abilities (Ferguson, 2003; Ford, 2013). Such results have now been found to be the case only because such measures are composed of White-specific culturally laden items. Tests are written from the cultural understanding of the ruling class without context for group cultures, which speak the same language, but have transformed it within their

contexts. This could cause meaningful shifts in words. Further there is the perception among educators that in the context of IQ performance, Blacks tend to show worse outcomes than Whites. However, numerous studies have found that it is not the case (Richardson, 2000, 2002; Nijenhuis et al., 2004; Serpell et al., 2006). Williams and Rivers (1972) showed that test instructions in Standard English penalize Black students, and that if the language of the test is put in familiar labels, without training or coaching, their performances on the tests increase significantly. If similar tests were written in Black-relevant language, or even street-smart contexts, a gifted individual in the White culture would underperform being in uncharted waters. As such researchers, for example, have called into question the use of IQ testing with Black populations (Kwate, 2001; Obiakor and Utley, 2004) positing that their specific cultural context is not fitted by the models used to build such assessments, and contributed historically to the mishandling of diagnostic and remediation in these populations. These tests must be constructed with culturally relevant contents and backgrounds to avoid misdiagnosis. For a strong assessment system, teachers should have knowledge of formal and informal measures of reading proficiency, and language dynamics, and be skilled in the use of these measures. Thus again, a way to mitigate the problem of language in learning, teaching, and developing thinking is to grasp a cultural understanding with teachers being trained to incorporate these elements in the classroom, and in their assessment tools. To conclude, the understanding of how culture inhibit one's ability to think will better help teachers to create tools and pedagogies to develop pathways to overcome language incomprehension-born stereotypes and biases, and avoid the mis- and underdiagnoses of Black individuals.

## TEACHERS MUST LEARN THE LANGUAGE

Generalized spread of biases have been found amongst educators when it comes to Black students. The dynamics at play in Black students' performance have been extensively researched. But what is the influence of language differences between educators and Black students? How much of variation in their academic performance can be explained by the lack of understanding of language specificities? Can language barrier contribute to the students underperforming? If language is essential to human learning, then practitioners need to do their best to learn the language of their students. This means relating classroom concepts to the experience of their students (Crogman et al., 2015). This could be one of the reasons why Black students tend to perform on average worse than other groups. As mentioned above black students underperformed when a stereotype was attached to the task they had to accomplish (Stone et al., 1999), the question is, how much does the lack of relatedness and understanding of language and cultural identity affect the perception of stereotype, and thereby the quality of academic performance?

American instructional structures have been based from their origin on principles, concepts, and languages formed and understood by the ruling class, the *White* culture. A recent study (Gilliam et al., 2016) found that racial bias in relationships

between teachers and students goes as far back as preschool. The study showed that Black children are most negatively impacted. In this context, children are generally taught to think in a language outside of their real-life experience, and how could it be any other way when teachers are the one person teaching them, and the unique channel to their academic education? The safest way would be one to one teaching or segregated classrooms with the context of the teacher and the student being similar. This model has historically been rejected. Crogman et al. (2015) showed that the difference for the success in the Finnish school system is due to its singular structure. The American school system is much more culturally multi-faceted, with a deep-rooted history of racial prejudice, where racial biases are prebuilt into the language of instruction.

The understanding of the language that is part of our experience affects what happens in the classroom and the ways in which learners begin to understand the relationship between their own language and that of their learning. What happens when students and teachers struggle to bridge the cultural gaps that exist between them, and their relationships suffer as a result? What can be done to change it? Hernández et al. (2016) reported that verbal competence indirectly predicted higher academic adjustment *via* lower teacher-student conflict. Spilt and Hughes (2015) showed that Black ethnicity, and not IQ and SES, uniquely predicted atypical conflict trajectories while controlling for sociobehavioral predictors. Black children were at risk of increasingly conflicted relationships with elementary school teachers, which has been found to increase the risk of academic underachievement in middle school (Spilt and Hughes, 2015).

An important question in the case of Black children is, are these biases eliminated when the teachers themselves are in this group? The short answer is—not really. The downside however is that because classroom preparation, books, exercise, teacher training are closely related to the language experience of the dominant group, Black teachers are not any more effective than their Caucasian counter parts (Garcia and Guerra, 2004). As such, it would seem wise to recruit teachers from areas, which have the same experience and language than the children there to carry out instruction. Nonetheless, since this might somewhat be an impossible task, then all teachers teaching in such neighborhoods should get to know these communities in their both languages and experiences. If the language pedagogies focus on the interpretation and creation of meaning, language is learned as a system of personal engagement with a new world, where learners engage with diversity at a personal level. This will help them examine their own cultural biases and beliefs. The same argument is cited for effective policing (Pickett, 2007).

Ideally a child's language development should be evaluated in terms of his/her progress toward the norms for his/her own particular speech community (Cadzen, 1966). The Black-White achievement gap is examined through the following factors: teacher quality, academic rigor, high academic expectations, family involvement, and exposure to literacy-enriched environments. These significantly influence students' achievement (Van Kleeck 2004; Wasik and Hendrickson, 2004; Barton and Coley, 2009; Edwards and Turner, 2009). Research has overlooked the role of



language proficiency and culture in teachers and students. It is true that family involvement is crucial, but it is not clear how the other factors get around the existing implicit biases without cultivating a pedagogy that is sensitive to the background of the learners. It is well documented that the Black–White achievement gap has continued to widen since the 1980s. Ferguson (2003) reminded us that teachers' perceptions, expectations, and behaviors interact with students' beliefs, behaviors, and work habits in ways that help to perpetuate the Black–White test score gap for example.

The achievement gap will be normalized when classrooms are more streamlined to cultural norms. The environment in which students interact is paramount. The environment in which the students learn must be perceived as safe and relatable. Educators must understand how culture and language are constructed for the various groups in the class—i.e., the teacher must create a place that is based on, and promotes cultural understanding (Crogman et al., 2015). Such an environment will allow students to ask questions and develop critical thinking skills free of constraints unique to language barriers. Crogman and Trebeau Crogman (2016) have demonstrated the vital influence of question-asking on student learning. **Figure 1** explores this concept, showing the emergence of question asking from curiosity to inquiry. Question-asking strategies, carefully developed in instruction, is one way to remedy positively to language deficits in student learning.

## CONCLUSION

This essay is responding in some context to the language-thought hypothesis, which suggests that there is no evidence that language influences thoughts. The thinking process is refined by the questions we ask and this is clearly demonstrated in the literature (Crogman and Trebeau Crogman, 2016). There is large body of evidence that shows that language helps us to formulate better questions, which is essential for the critical thinking process.

## 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
- Aloqaili, A. S. (2012). The relationship between reading comprehension and critical thinking: a theoretical study. *J. King Saud Univ. Lang. Transl.* 24, 35–41. doi:10.1016/j.jksult.2011.01.001
- American Heritage Dictionary. (1983). *Bias*. Boston: Houghton Mifflin.
- Baker, L. (2009). Metacognitive processes and reading comprehension In: Israel SE, Duffy GG, Beall LC, eds. *Handbook of Research on Reading Comprehension*, New York, NY: Routledge, 373–378.
- Barton, P. E., and Coley, R. J. (2009). *Parsing the Achievement Gap II. Policy Information Report*. Educational Testing Service. Princeton, NJ. Available at: <https://www.ets.org/Media/Research/pdf/PICPARSINGII.pdf>
- Berlyne, D. E. (1954). A theory of human curiosity. *Br. J. Psychol.* 45, 180–191.
- Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill.
- Berlyne, D. E. (1965). *Structure and Direction in Thinking*. New York, NY: Wiley.
- Berlyne, D. E. (1966). Curiosity and exploration. *Science* 153, 25–33.1. doi:10.1126/science.153.3731.25
- Bigler, R. S., and Liben, L. S. (2007). Developmental intergroup theory: explaining and reducing children's social stereotyping and prejudice. *Curr. Dir. Psychol. Sci.* 16, 162–166. doi:10.1111/j.1467-8721.2007.00496.x
- Bloom, B., Englehart, M., Furst, E., Hill, W., and Krathwohl, D. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. New York, NY: Longmans Green.
- Bloom, P., and Keil, F. C. (2001). Thinking through language. *Mind Lang.* 16, 351–367. doi:10.1111/1468-0017.00175
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cogn. Psychol.* 43, 1–22. doi:10.1006/cogp.2001.0748
- Boroditsky, L. (2003). "Linguistic relativity," in *Encyclopedia of Cognitive Science*, ed. L. Nadel (London: MacMillan), 917–921.
- Borowske, K. (2005). "Curiosity and motivation-to-learn," in *Comunicación Presentada a la ACRL 12th National Conference* (Minneapolis, MN).
- Cadzen, C. B. (1966). Subcultural differences in child language: an interdisciplinary review. *Merrill Palmer Q.* 1955, 185–214.
- Catts, H. W., Hogan, T. P., and Adlof, S. M. (2005). "Developmental changes in reading and reading disabilities," in *The Connections between Language and Reading Disabilities*, eds H. W. Catts and A. G. Kamhi (Mahwah, NJ: Lawrence Erlbaum Associates), 23–36.
- Catts, H. W., and Kamhi, A. G. (2005). *Language and Reading Disabilities*, 2nd Edn. Boston: Pearson.
- Cheney, D. L., and Seyfarth, R. M. (1998). "Why animals don't have language," in *Tanner Lectures on Human Values*, Vol. 19, 173–210.
- Cherng, H. Y. S. (2017). If they think I can: teacher bias and youth of color expectations and achievement. *Soc. Sci. Res.* 66, 170–186. doi:10.1016/j.ssresearch.2017.04.001

Further, humans achieve much more than other animals because of language, so much so that when language is lost, human mental faculties are impaired. Additionally, it is a fair conclusion that language is probably the most important domain in the development of critical thinking skills. Thus researchers must consider ways in which language should be reconsidered as an important tenet of learning models, and the role it plays to influence all other domains in their learning theories.

We cannot escape the fact that our failure to grapple with the impact of language in its cultural norms, have caused a tremendous burden for learning and instruction in the classroom. Ford (2013) recommended a greater reliance on performance-based assessments and non-verbal intelligence tests. Non-verbal measures reduce the reliance on language and social-cultural influences. A question that we must think on is: is the emphasis on reading the only requirement to be functional in society? This question will be best answered when we stop classification of dyslexia as a problem. As we learn to appreciate culture in a more positive way, and the role it plays in the formulation of language, we will be able to create curricula that are beneficial to all groups. In this essay we modified the learning model of Crogman and Trebeau Crogman (2016) by adding the need to consider the role of language and comprehension, and how they impact the other domains. Further language is proposed as fundamental to the question process. Understanding the role of language in the student's reference frame will help educators guide them, and students develop their critical thinking skills *via* better question asking processes, while overcoming associated learning deficits. The role of the instructor is pivotal for success. As Cherng (2017) sees it, the solution to this dilemma is in instructors' better training and awareness to put an end to the implicit biases in education.

## AUTHOR CONTRIBUTIONS

All work was done by HC.

- Cherng, H. Y. S., and Halpin, P. F. (2016). The importance of minority teachers: student perceptions of minority versus white teachers. *Educ Res* 45, 407–420. doi:10.3102/0013189X16671718
- Chomsky, N. (1956). “Three models for the description of languages,” in *I. R. E. Transactions on Information Theory IT*, 1956.
- Chomsky, N. A. (1975). *Reflections on Language*. New York: Pantheon Books.
- Clark, K. B., and Clark, M. K. (1939). The development of consciousness of self and the *psychology*. *J. Soc. Psychol.* 10, 591–599. doi:10.1080/00224545.1939.9713394
- Conley, D. T. (2007b). *Redefining College Readiness (Volume 3)*. Eugene, OR: Educational Policy Improvement Center.
- Cowie, F. (1997). The logical problem of language acquisition. *Synthese* 111, 17–51. doi:10.1023/A:1004975305820
- Crogman, H., and Trebeau Crogman, M. (2016). Generated questions learning model (GQLM): beyond learning styles. *Cogent Educ.* 3, 1202460. doi:10.1080/2331186X.2016.1202460
- Crogman, T. H., Trebeau Crogman, A. M., Warner, L., Mustafa, A., and Peters, R. (2015). Developing a new teaching paradigm for the 21st century learners in the context of Socratic methodologies. *Br. J. Educ. Soc. Behav. Sci.* 9, 62–95. doi:10.9734/BJESBS/2015/17825
- Davidson, B. W. (1994). Obstacles and opportunities for critical thinking in Japan. *Paper Presented at the 14th Annual Meeting of the International Conference on Critical Thinking and Educational Reform*, Sonoma, CA. ERIC database. (ED377216).
- Davis, K. (2009). “A girl like me,” in *In Struck by lightning: The Transdiasporan Phenomenon of Skin Bleaching JENdA: A Journal of Culture and African Women Studies*, 14. Available at: <https://www.africaknowledgeproject.org/index.php/jenda/issue/view/41>
- de Valenzuela, J. S., Copeland, S. R., Qi, C. H., and Park, M. (2006). Examining educational equity: revisiting the disproportionate representation of minority students in special education. *Except. Child.* 72, 425–441. doi:10.1177/001440290607200403
- Devitt, M. (2006). *Ignorance of Language*. Oxford: Oxford University Press.
- Diamond, A. (2006). “The early development of executive functions,” in *Lifespan Cognition: Mechanisms of Change*, Chap. 6, eds E. Bialystok and F. I. M. Craik (New York: Oxford University Press), 70–95.
- DuBois, W. E. B. (1903). *The Negro Church*. Atlanta: Atlanta University Press, 1.
- Edwards, H. (1973). *Sociology of Sport*. Homewood, IL: Dorsey Press.
- Edwards, P. A., and Turner, J. D. (2009). “Family literacy and reading comprehension,” in *Handbook of Research on Reading Comprehension*, eds S. E. Israel and G. G. Duffy (New York, NY: Routledge), 622–642.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1996). *Rethinking Innateness*. Cambridge, MA: MIT Press.
- Ennis, R. H. (1998). Is critical thinking culturally biased? *Teach. Philos.* 21, 15–33. doi:10.5840/teachphil19982113
- Ferguson, R. F. (2003). Teachers’ perceptions and expectations and the black-white test score gap. *Urban Educ.* 38, 460–507. doi:10.1177/0042085903038004006
- Fernyhough, C. (2008). *A Thousand Days of Wonder: A Scientist’s Chronicle of His Daughter’s Developing Mind*. New York, NY: Penguin.
- Fleming, J. (1984). *Blacks in College. A Comparative Study of Students’ Success in Black and in White Institutions*. San Francisco, CA: Jossey-Bass, 94194.
- Ford, D. Y. (2010). *Reversing Underachievement among Gifted Black Students*, 2nd Edn. Waco, TX: Prufrock Press.
- Ford, D. Y. (2013). *Recruiting and Retaining Culturally Different Students in Gifted Education*. Waco, TX: Prufrock Press.
- Garcia, S. B., and Guerra, P. L. (2004). Deconstructing deficit thinking: working with educators to create more equitable learning environments. *Educ. Urban Soc.* 36, 150–168. doi:10.1177/0013124503261322
- Gardner, R. III, and Hsin, Y. (2008). “Reading and children with emotional/behavioral disorders,” in *Balanced Reading Instruction*, 25–40.
- Gernsbacher, M. A. (1990). *Language Comprehension as Structure Building*. Hillsdale, NJ: Erlbaum.
- Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., and Shic, F. (2016). *Do Early Educators’ Implicit Biases Regarding Sex and Race Relate to Behavior Expectations and Recommendations of Preschool Expulsions and Suspensions?* New Haven: Yale Child Study Center, 991–1013.
- Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics* 30, 5–55.
- Green, L. (1999). “Stereotypes: negative racial stereotypes and their effect on attitudes toward African-Americans,” in *Perspectives on Multiculturalism and Cultural Identity*, Vol. 9(1). Available at: <http://www.ferris.edu/jimcrow/links/VCU.htm>
- Green, L. J. (2002). *African American English: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Hadley, O. A. (2000). *Teaching Language in Context*, 3rd Edn. Boston: Heinle: Heinle.
- Halliday, M. A. (1993). Towards a language-based theory of learning. *Linguist. Educ.* 5, 93–116. doi:10.1016/0898-5898(93)90026-7
- Hamilton, D. L., and Trolier, T. K. (1986). *Stereotypes and Stereotyping: An Overview of the Cognitive Approach*. San Diego, CA: Academic Press.
- Hardin, C., and Banaji, M. R. (1993). The influence of language on thought. *Soc. Cogn.* 11, 277–308. doi:10.1521/soco.1993.11.3.277
- Harris, C. L. (2006). “Language and cognition,” in *Encyclopedia of Cognitive Science*. John Wiley & Sons.
- Hermer, L., and Spelke, E. S. (1996). Modularity and development: the case of spatial reorientation. *Cognition* 61, 195–232. doi:10.1016/S0010-0277(96)00714-7
- Hermer-Vasquez, L., Spelke, E. S., and Katsnelson, A. S. (1999). Sources of flexibility in human cognition: dual-task studies of space and language. *Cogn. Psychol.* 39, 3–36. doi:10.1006/cogp.1998.0713
- Hermer-Vasquez, L., Moffet, A., and Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: the case of two spatial memory tasks. *Cognition* 79, 263–299. doi:10.1016/S0010-0277(00)00120-7
- Hernández, M. M., Eisenberg, N., Valiente, C., VanSchyndel, S. K., Spinrad, T. L., Silva, K. M., et al. (2016). Emotional expression in school context, social relationships, and academic adjustment in kindergarten. *Emotion* 16, 553. doi:10.1037/emo0000147
- Hoyles, A., and Hoyles, M. (2010). Race and dyslexia. *Race Ethn. Educ.* 13, 209–231. doi:10.1080/13613321003726835
- Hudson, R. F., High, L., and Al Otaiba, S. (2007). Dyslexia and the brain: what does current research tell us? *Read. Teach.* 60, 506–515. doi:10.1598/RT.60.6.1
- Hurtado, S., Cabrera, N. L., Lin, M. H., Arellano, L., and Espinosa, L. L. (2009). Diversifying science: underrepresented student experiences in structured research programs. *Res High Educ* 50, 189–214. doi:10.1007/s11162-008-9114-7
- Hutt, C., and Bhavnani, R. (1972). Predictions from play. *Nature* 237, 171–172. doi:10.1038/237171b0
- Inhelder, B. (1978). Language and thought: some remarks on Chomsky and Piaget. *J. Psycholinguist. Res.* 7, 263–268. doi:10.1007/BF01068109
- Jencks, C., and Phillips, M. (eds) (1998). *The Black-White Test Score Gap*. Washington D.C.: Brookings Institution Press.
- Jewell, S. K. (1993). *From Mammy to Miss America and Beyond: Cultural Images and the Shaping of US Policy*. New York: Routledge.
- Kalia, A. A., Legge, G. E., and Giudice, N. A. (2008). Learning building layouts with non-geometric visual information: the effects of visual impairment and age. *Perception* 37, 1677–1699.1. doi:10.1068/p5915
- Kandeou, P., Rapp, D. N., and Van Den Broek, P. (2003). *The Influence of Readers’ Prior Knowledge on Text Comprehension and Learning from Text. Progress in Education*, Vol. 13. New York: Nova Science Publishers, Inc, 189–209.
- Kay, P., and Kempton, W. (1984). What is the Sapir-Whorf-hypothesis? *Am. Anthropol.* 86, 65–79. doi:10.1525/aa.1984.86.1.02a00050
- Kubota, R. (2004). “Critical multiculturalism and second language education,” in *Critical Pedagogies and Language Learning*, eds B. Norton and K. Toohey (Cambridge: Cambridge University Press), 30–52.
- Kubota, R. (2013). “Multiculturalism and second language learning,” in *Encyclopedia of Applied Linguistics*. Available at: <https://www.scribd.com/document/339141896/The-Encyclopedia-of-Applied-Linguistics>
- Kwate, N. (2001). Intelligence or misorientation? Eurocentrism in the WISC-III. *J. Black Psychol.* 27, 221–238. doi:10.1177/0095798401027002005
- Lemke, J. L. (1990). *Talking Science: Language, Learning, and Values*. Norwood, NJ: Ablex.
- Levinson, S. C., and Wilkins, D. P. (eds) (2006). *Grammars of Space: Explorations in Cognitive Diversity*, Vol. 6. New York: Cambridge University Press.
- Levy, S. R., and Dweck, C. S. (1999). The impact of children’s static versus dynamic conceptions of people on stereotype formation. *Child Dev.* 70, 1163–1180. doi:10.1111/1467-8624.00085
- Lindo, E. J. (2006). The African American presence in reading intervention research. *Remed. Spec. Educ.* 27, 148–153. doi:10.1177/07419325060270030301

- Locke, J. (1975). *An Essay Concerning Human Understanding*, ed. P. Nidditch (Oxford: Clarendon).
- Locks, A. M., Hurtado, S., Bowman, N. A., and Oseguera, L. (2008). Extending notions of campus climate and diversity to students' transition to college. *Rev. High. Educ.* Vol. 31, 257–285.
- Loewenstein, G. (1994). The psychology of curiosity: a review and reinterpretation. *Psychol. Bull.* 116, 75–98. doi:10.1037/0033-2909.116.1.75
- Lyons, G. R., Shaywitz, S. E., and Shaywitz, B. A. (2003). A definition of dyslexia. *Ann. Dyslexia* 53, 1–14. doi:10.1007/s11881-003-0001-9
- Majid, A., and Levinson, S. C. (2011). The senses in language and culture. *Senses Soc.* 6, 5–18. doi:10.2752/174589311X12893982233551
- Malone, T. W. (1980). "What makes things fun to learn? Heuristics for designing instructional computer games," in *Proceedings of the 3rd ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems* (New York, NY: ACM), 162–169.
- Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cogn. Sci.* 5, 333–369. doi:10.1207/s15516709cog0504\_2
- Markman, E., and Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cogn. Psychol.* 16, 1–27. doi:10.1016/0010-0285(84)90002-1
- Martin, J. R. (1991). "Nominalization in science and humanities: distilling knowledge and scaffolding text," in *Functional and Systemic Linguistics: Approaches and Uses*, ed. E. Ventola (Berlin: de Gruyter), 307–337. (Trends in linguistics studies and monographs 55).
- Martin, N. M., and Duke, N. K. (2011). "Intervention to enhance informational text comprehension," in *Handbook of Reading Disability Research*, eds A. McGill-Franzen and R. L. Allington (New York, NY: Routledge), 345–361.
- McKown, C., and Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in middle childhood. *Child Dev.* 74, 498–515. doi:10.1111/1467-8624.7402012
- Nida, E. (1998). Language, culture and translation. in *Foreign Lang. J.* 115, 29–33.
- Nijenhuis, J. T., Resing, W., Tolboom, E., and Bleichrodt, N. (2004). Short-term memory as an additional predictor of school achievement for immigrant children? *Intelligence* 32, 203–213. doi:10.1016/j.intell.2003.07.001
- Obiakor, F. E., and Utley, C. A. (2004). Educating culturally diverse learners with exceptionalities: a critical analysis of the brown case. *Peabody J. Educ.* 79, 141–156. doi:10.1207/s15327930pje7902\_10
- Paul, R. W., and Adamson, K. R. (1990). Using critical thinking to identify bias in the news media: the art of critical analysis. *Soc. Stud. Rev.* 29, 11–28.
- Peffley, M., Hurwitz, J., and Sniderman, P. M. (1997). Racial stereotypes and whites' political views of blacks in the context of welfare and crime. *Am. J. Pol. Sci.* 41(1):30–60. doi:10.2307/2111708
- Pelham, B. W., Mirenberg, M. C., and Jones, J. T. (2002). Why Susie sells seashells by the seashore: implicit egotism and major life decisions. *J. Pers. Soc. Psychol.* 82, 469. doi:10.1037/0022-3514.82.4.469
- Piaget, J. (1976). *Piaget Sampler: An Introduction of Jean Piaget through His Own Words*. New York: Wiley.
- Pica, P., Lemer, C., Izard, V., and Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499–503. doi:10.1126/science.1102085
- Pickett, X. (2007). Policing black communities. *Pub. Justice* 30. Available at: [https://www.cpjjustice.org/uploads/Policing\\_Black\\_Communities.pdf](https://www.cpjjustice.org/uploads/Policing_Black_Communities.pdf)
- Plous, S., and Williams, T. (1995). Racial stereotypes from the days of American slavery: a continuing legacy. *J. Appl. Soc. Psychol.* 25, 795–817. doi:10.1111/j.1559-1816.1995.tb01776.x
- Proctor, S. L., Graves, S. L., and Esch, R. C. (2012). Assessing African American students for specific learning disabilities: the promises and perils of response of intervention. *J. Negro Educ.* 81, 268–282. doi:10.7709/jnegroeducation.81.3.0268
- Ricci, F. A., and Huang, H. F. (2013). Cultural conditioning: influences on critical thinking. *J. Worldwide Forum Educ. Cult.* 5, 47.
- Richardson, K. (2000). *The Making of Intelligence*. New York: Columbia University Press.
- Richardson, K. (2002). What IQ tests test. *Theory Psychol.* 12, 283. doi:10.1177/0959354302012003012
- Robinson, S. A. (2013). Educating black males with dyslexia. *Interdiscip. J. Teach. Learn.* 3, 159–174.
- Rogoff, B. E., and Lave, J. E. (1984). *Everyday Cognition: Its Development in Social Context*. Cambridge, MA: Harvard University Press.
- Romano, A. R. (2007). *Elements of a Context-Specific Model of Critical Thinking in Management: A Delphi Study*. ProQuest.
- Schaller, S. (1991). *A Man Without Words*. Berkeley, CA: University of California Press.
- Serpell, Z. N., Boykin, A. W., Madhere, S., and Nasim, A. (2006). The significance of contextual factors in African American students' transfer of learning. *J. Black Psychol.* 32, 418–441. doi:10.1177/0095798406292466
- Sinanović, O., Mrkonjić, Z., Zukić, S., Vidović, M., and Imamović, K. (2011). Post-stroke language disorders. *Acta Clin. Croat.* 50, 79–93.
- Smedley, B. D., Myers, H. F., and Harrell, S. P. (1993). Minority-status stresses and the college adjustment of ethnic minority freshmen. *J. Higher Educ.* 64, 434–452. doi:10.1080/00221546.1993.11778438
- Snow, C. E., Burns, M. S., and Griffin, P. (1998). *Preventing Reading Difficulties in Young Children*. Washington, DC: National Academy Press.
- Sparks, J. R. (2012). "Language/discourse comprehension and understanding," in *Encyclopedia of the Sciences of Learning*. US: Springer, 1713–1717.
- Spelke, E. (2010). "Innateness, choice, and language," in *Chomsky Notebook*, eds J. Bricmont and J. Franck (New York, NY: Columbia University Press), 203–210.
- Spelke, E. S., and Hermer, L. (1996). in *Early Cognitive Development: Objects and Space, Perceptual and Cognitive Development*, eds R. Gelman and T. Kit-Fong (San Diego: Academic Press), 71–114
- Spilt, J. L., and Hughes, J. N. (2015). African American children at risk of increasingly conflicted teacher–student relationships in elementary school. *School Psych. Rev.* 44, 306–314. doi:10.17105/spr-14-0033.1
- Stone, J., Lynch, C. I., Sjomeling, M., and Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *J. Pers. Soc. Psychol.* 77, 1213. doi:10.1037/0022-3514.77.6.1213
- Tversky, A., and Kahneman, D. (1985). "The framing of decisions and the psychology of choice," in *Environmental Impact Assessment, Technology Assessment, and Risk Analysis: Contributions from the Psychological and Decision Sciences*, eds V. T. Covello, J. L. Mumpower, P. J. M. Stallen, and V. R. R. Uppuluri (Berlin, Heidelberg: Springer), 107–129.
- van Kleeck, A. (2004). "Fostering preliteracy development via storybook sharing interactions: The cultural context of mainstream family practices," in *Handbook of Language and Literacy*, eds C. Stone, E. Silliman, B. Ehren, and K. Apel (New York: Guilford), 175–208.
- Vellutino, F. R., and Fletcher, J. M. (2005). "Developmental dyslexia," in *The Science of Reading: A Handbook*, eds M. J. Snowling and C. Hulme (Malden, MA: Blackwell, Publishing), 362–378.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Wasik, B. H., and Hendrickson, J. S. (2004). "Family literacy practices," in *Handbook of Language and Literacy: Development and Disorders*, eds C. A. Stone, E. R. Silliman, B. J. Ehren, and K. Apel (New York, NY: Guilford Press), 154–174.
- Waxman, S., and Kosowski, T. (1990). Nouns mark category relations: Toddlers' and Preschoolers' word-learning biases. *Child Dev.* 61, 1461–1473. doi:10.2307/1130756
- Weiler, N. (2015). Speaking a second language may change how you see the world. *Sci. Mag.* Available at: <http://www.sciencemag.org/news/2015/03/speaking-second-language-may-change-how-you-see-world>
- Welch, K. (2007). Black criminal stereotypes and racial profiling. *J. Contemp. Crim. Justice* 23, 276–288. doi:10.1177/1043986207306870
- West-Olatuji, C. A., Baker, J. C., and Brooks, M. (2006). African American adolescents males: giving voice to their educational experiences. *Multicultural Perspect.* 8, 3–9. doi:10.1207/s15327892mcp0804\_2
- West's Encyclopedia of American Law. (2005). *Bias*. Detroit: Thomson and Gale.
- Whiting, G. (2009). Gifted black males: understanding and decreasing barriers to achievement and identity. *Roeper Rev.* 31, 224–233. doi:10.1080/02783190903177598
- Whorf, B. L. (1956). *Language, Thought and Reality*. Cambridge, MA: MIT Press.
- Williams, R. L., and Rivers, L. W. (1972). The use of standard and nonstandard English in testing black children. In *annual meeting of the American Psychological Association, Hawaii, September*.
- Zhao, N. (2008). A critical reading of multiculturalism in English language teaching and learning. *Online Submission* 5, 55–60.



Zlatev, J., and Blomberg, J. (2015). Language may indeed influence thought. *Front. Psychol.* 6:1631. doi:10.3389/fpsyg.2015.01631

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Crogman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Screening Protocol for Early Identification of Brazilian Children at Risk for Dyslexia

Giseli D. Germano<sup>1,2\*</sup>, Alexandra B. P. de C. César<sup>1,3</sup> and Simone A. Capellini<sup>1</sup>

<sup>1</sup> Investigation Learning Disabilities Laboratory, Department of Speech and Hearing Sciences, São Paulo State University "Júlio de Mesquita Filho" (UNESP), Marília, Brazil, <sup>2</sup> Department of Special Education, São Paulo State University "Júlio de Mesquita Filho" (UNESP), Marília, Brazil, <sup>3</sup> Speech and Hearing Sciences Department, São Paulo State University "Júlio de Mesquita Filho" (UNESP), Marília, Brazil

## OPEN ACCESS

### Edited by:

Layne Kalbfleisch,  
George Washington University,  
United States

### Reviewed by:

Angela Jocelyn Fawcett,  
Swansea University, United Kingdom  
Pin-Ju Chen,  
St. Mary's Junior College of Medicine,  
Nursing and Management, Taiwan

### \*Correspondence:

Giseli D. Germano  
giseliger@yahoo.com.br

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

Received: 06 April 2017

Accepted: 22 September 2017

Published: 27 October 2017

### Citation:

Germano GD, César ABPC and  
Capellini SA (2017) Screening  
Protocol for Early Identification of  
Brazilian Children at Risk for Dyslexia.  
Front. Psychol. 8:1763.  
doi: 10.3389/fpsyg.2017.01763

Early identification of students at risk of dyslexia has been an educational challenge in the past years. This research had two main goals. First, we aimed to develop a screening protocol for early identification of Brazilian children at risk for dyslexia; second, we aimed to identify the predictive variables of this protocol using Principal Component Analysis. The major step involved in developing this protocol was the selection of variables, which were chosen based on the literature review and linguistic criteria. The screening protocol was composed of seven cognitive-linguistic skills: Letter naming; Phonological Awareness (which comprises the following subtests: Rhyme production, Rhyme identification, Syllabic segmentation, Production of words from a given phoneme, Phonemic Synthesis, and Phonemic analysis); Phonological Working memory, Rapid naming Speed; Silent reading; Reading of words and non-words; and Auditory Comprehension of sentences from pictures. A total of 149 children, aged from 6 years to 6 and 11, of both genders who were enrolled in the 1st grade of elementary public schools were submitted to the screening protocol. Principal Component Analysis revealed four factors, accounting for 64.45% of the variance of the Protocol variables: first factor ("pre-reading"), second factor ("decoding"), third factor ("Reading"), and fourth factor "Auditory processing." The factors found corroborate those reported in the National and International literature and have been described as early signs of dyslexia and reading problems.

**Keywords:** reading, dyslexia, early identification, phonological awareness, assessment

## INTRODUCTION

Early identification of students at risk for dyslexia has been an educational challenge in the past years. Although scientific research has explored the nature, etiology, assessment, and intervention of this learning disorder, educators are still having a hard time recognizing its signs, which suggest that a child might be at risk for reading failure without being identified. Such early identification should allow interventions to be implemented before a downward spiral of underachievement, lowered self-esteem and poor motivation sets in (Shaywitz and Shaywitz, 2005; Kirby et al., 2010; Snowling, 2013; Hulme et al., 2015). In Brazil, this topic is still fairly new; research has been carried out since 2009 (Capellini et al., 2009, 2015; Andrade et al., 2011; Fadini and Capellini, 2011;

Fukuda and Capellini, 2011, 2012), attempting to develop a screening protocol for early identification of children at risk for dyslexia. These studies have reported that phonological awareness, verbal working memory, and rapid naming correspond to the central phonological mechanisms of acquiring reading and writing that have been also reported by De Jong and Van der Leij (1999). However, none of them explored the predictive values of each variable and their impacts on the development of a minimal protocol for early identification. Hulme and Snowling (2014) have considered letter knowledge, phonological awareness, and rapid automatized naming as predictors which are important, since it makes it possible to differentiate individual performance in students at risk for dyslexia, regarding decoding skills in alphabetic languages in the early stages. Studies have reported that students with developmental dyslexia may present as manifestations difficulties with accurate or fluent word recognition and spelling, even when they had received adequate instruction, and have no signs of fails in intelligence and sensory abilities (Shaywitz and Shaywitz, 2005; Kirby et al., 2010; Snowling, 2013; Hulme et al., 2015). The authors also described that dyslexia is the result of several risk factors, and children who have language difficulties in the first school years are usually considered as being at high risk for learning disabilities. Another important issue about dyslexia refers about family history, which also plays an important role as a predictor of literacy outcome in the preschool years. However, assessment's protocols will only help to identify the risk after the children start literacy at school, when they will have formal instruction about letter knowledge, phonological awareness, and rapid automatized naming (RAN); together these skills provide good sensitivity and specificity as a screening battery. Furthermore, the consensus between these studies is that the first signs of dyslexia include delays in speech and language development, with phonological memory (non-word repetition) and expressive language (naming) skills being particularly affected, as mentioned in the studies of Carroll et al. (2014) and Thompson et al. (2015).

The relationship between phonological awareness, rapid naming, and reading in alphabetic languages has been documented in the literature over the last decades. Germano et al. (2014) described that Brazilian Portuguese language has an alphabetic system and that most words can be successfully read through phonological decoding, according to grapheme-phoneme correspondences (Pinheiro et al., 2008). Scliar-Cabral (2003) described that reading in Brazilian Portuguese is considerate to be transparent since it presents a set of one-to-one graph phonological relations, that is, univocal relations, and a set of inconsistent relations, many of which are governed by rules. Thus, a characteristic of reading processing in Brazilian Portuguese is that it can be performed, almost successfully, only when the grapheme-phoneme matching rules are known and phonological decoding is used, mostly at the beginning of literacy acquisition. With regard to writing, spelling of Brazilian Portuguese is considered as being more opaque. The reason is because, in general, writing is considered a more complex cognitive process that requires intention, selection, planning, monitoring, and revision, as well as a specific coding process

(Godoy and Pinheiro, 2013). Therefore, learning to read and write implies a deliberate reflection of speech, promoting metalinguistic awareness (Bradley and Bryant, 1983; Hayes and Slater, 2008; Manz et al., 2010). Learning to read requires highly complex task, such as visual integration, orthographic, phonological, and semantic information. For example, in the dual route model of reading aloud (Coltheart et al., 2001; Ziegler et al., 2008), the reading process also requires a series of interacting stages, from letter feature detection to phonological output processes. This process was divided into two major routes: the lexical orthographic route and the non-lexical phonological route. The lexical route is important because allows the correct pronunciation of irregular words, while the non-lexical route allows the pronunciation of novel words and non-words, using not only phonological processes but also letter perception. The reading circuit is composed of neural systems including phonology, morphology, syntax, and semantics, as well as other processes such as visual and orthographic processes, that requires memory, attention, comprehension, eye movements, and cognition. By having these skills, the reader develops the so-called automatism, reading with adequate precision and speed. When this process becomes automatic, the effort toward the act of reading becomes less apparent (Norton and Wolf, 2012). Despite the vast international literature on this theme, there is lack of research on these precursors in Brazilian Portuguese Language, concerning first graders. In Brazilian Portuguese language, most words can be successfully read using phonological decoding, and even reading fluency reflect the ability of the reader to use grapheme-phoneme correspondences (Pinheiro et al., 2008). In Brazilian Portuguese, beginner readers, from first to third graders predominates the use of phonological route, but with age, performance gradually relies lexical route, such as knowledge and sight word vocabulary (Oliveira and Capellini, 2010; Mota et al., 2012). Phonological awareness is one of the most important precursor skill of reading and spelling and also one important predictors of the word recognition difficulties that characterize developmental dyslexia, one of the most common learning disorders, as reported in Peterson and Pennington (2012) and Skeide et al. (2015). Phonological awareness is the ability to identify, distinguish, and manipulate sounds within spoken language, and its importance to reading is widely acknowledged; thus, children who are able to identify and manipulate individual sounds have good academic performance. Impairment to the availability of early phonological skills can hinder subsequent reading progress (Duncan et al., 2013).

Studies have reported that the development of phonological awareness skills occurs in a sequential pattern, beginning in the 1st months of a child's life, before entering school. However, these skills have been described to have an important role in reading acquisition, as the perception that speech has an underlying phonemic structure allows storage in long-term phonological memory, using the generative mechanism of phonological memory, which converts spellings into phonology (Chard and Dickson, 1999; Cervera-Mérida and Ygual-fernández, 2003; Gombert, 2003; Hayes and Slater, 2008; Germano and Capellini, 2011). Along with phonological awareness, phonological memory or verbal short-term memory (capacity of temporary storage

based on sound information) has been highlighted as a component of phonological processing, that is required in reading development (Wagner and Torgesen, 1987; Gathercole et al., 1999; Alloway et al., 2005). Thus, phonological memory has an important role for vocabulary acquisition, because it provides a temporary phonological representation of unfamiliar words, and later it will be responsible for an enduring representation in long-term memory (Gathercole and Baddeley, 1989; De Jong and Olson, 2004). It also contributes to the acquisition of letter knowledge (De Jong and Olson, 2004), facilitating the word identification when grapheme–phoneme correspondence rules is necessary, and facilitates text comprehension because allows children to recuperate words they have already read.

Hulme et al. (2015) have reported that the development of reading skills requires underlying ability of oral language abilities. Phonological skills has a causal influence on the later development of early word-level literacy skills, which has an impact in reading-comprehension, involving (semantic and syntactic) language skills. The authors presented a longitudinal study comparing children at familial risk for dyslexia, children with preschool language difficulties, and typically developing control children. Their findings described that as preschool measures of oral language it was found that phoneme awareness and grapheme-phoneme knowledge were important to acquire before school entry, which in turn predicted word-level literacy skills shortly after school entry. These results were indicated also for both typically developing children and those at risk of literacy difficulties. The authors highlighted the importance of oral language skills for the development of both word-level literacy and reading comprehension.

In addition, less speed in naming may reflect difficulty in the integration of cognitive and linguistic processes involved in fluent reading (Araújo et al., 2016). Studies (Jones et al., 2010; Araújo et al., 2016) using the Rapid Automatized Naming Test (RAN) (Denckla and Rudel, 1976), which was designed to measure the speed at which a series of highly familiar items such as letters, digits, objects, and colors can be named. As a cognitive requirement, visual naming represents a demanding array of attentional, perceptual, conceptual, memory, lexical, and articulatory processes. Wolf et al. (2000) argued that this, in turn, RAN has played an important rule for identification or recognition processes, which integrate information of present stimulus with known mental representations, quality that will influence the speed of processing. Lexical processes, that include semantic, phonological access and retrieval processes, can be integrated with cumulative information. After the cognitive processes, motor commands translate this phonological information into an articulated name. The entire process occurs within 500 ms. Difficulties have been found to be invariant across languages (Brizzolara et al., 2006; Capellini and Conrado, 2009; Araújo et al., 2010). One of the reasons for using naming speed as part of reading evaluations is because naming speed and reading are similar. According to Kirby et al. (2010), in both RAN and oral reading subjects are solicited to move their eyes sequentially across the page, encode the stimulus that they are focusing on, access the mental representation of that stimulus, and then activate the associated motor commands that allows the

subject to name that stimulus. Before the first motor commands is completed, the eyes must move on to the next stimulus, and so on. Just as in reading, the eyes must make a sweep back to the beginning of the next line. Several studies have justified the relationship between word reading and RAN concerning phonological deficits or phonological processing (Morris et al., 1998; Vaessen et al., 2009).

According to Thompson et al. (2015), identifying children with dyslexia or at risk for dyslexia means assessing the probability that a group of variables will identify positive cases of dyslexia (sensitivity), aiming to avoid false positives (specificity). Thus, the present study discusses the hypothesis that precursors of dyslexia described in Brazilian and international literature, such as knowledge of the alphabet, phonological awareness, working memory, rapid automatic naming, visual attention, reading words, and non-words, could be addressed for early identification of first grade children at risk for dyslexia in Brazil. This research had two main goals. First, we aimed to develop a screening protocol for early identification of children at risk for dyslexia; second, we aimed to identify the predictive variables of this protocol using Principal Component Analysis.

## MATERIALS AND METHODS

In the first part of this study, the steps to develop this screening protocol, such as variable selection, will be described based on the literature. This screening protocol was developed to be used as a universal screening for first grade children and as part of the Tier 1 of the response to intervention (RTI) model. According to Johnston and Kirby (2006), Tier 1 aims to identify the risks for behavioral and learning problems using procedures based on the academic curriculum of these children; therefore, it would be possible to verify if these children reached the expected results at their grade level. Capellini et al. (2015) used the Screening Protocol for the Early Identification of Reading Problems in Brazilian children at risk for dyslexia as part of a RTI study. Of the 156 students that were evaluated by these authors using the protocol, 62 fulfilled the risk criteria (performance below the 25th percentile for at least 51% of the Protocol variables). The students were submitted to phonological intervention, and the results obtained in the post-tests indicated that 12 students continued to be at risk, according to their performance. These students underwent a multidisciplinary evaluation to confirm the diagnosis.

Because reading involves multiple linguistic, visual, and attentional processes, it is likely that variable patterns of weaknesses may contribute to reading difficulties among children, as mentioned by Norton and Wolf (2012). However, the present study considered the recent investigations that have demonstrated that dyslexic children may have difficulties in underlying processes (e.g., phonological awareness and rapid naming test) and difficulties with RAN, related to visual attention processing (Franceschini et al., 2012; Germano et al., 2014). Taking that into consideration, the development of the screening protocol for early identification of reading problems (Capellini et al., 2009, 2015, 2017) was based on a literature review to

identify the skills for effective reading and writing. The Protocol was composed of seven cognitive-linguistic skills divided into seven tests. The tests and justification for their selection are shown in **Table 1**.

After selecting the tests, the next step concerned the choice of linguistic stimuli to compose the Protocol. This study was based on a phonological perspective called linear model and on the hierarchical model (Câmera, 1970a,b; Selkirk, 1982). The screening protocol was composed of words from a word bank created for this study; these words were extracted from 1th to 5th grade textbooks (elementary school) written in Portuguese (Germano and Capellini, 2008; Germano, 2011). This word bank included words belonging to different word classes or parts of speech, such as pronouns, prepositions, adjectives, adverbs, verbs, and nouns. Exclusion criteria were as follows: pronouns, prepositions, words that could vary according to the class or grammatical category, gender, and agreement, which happens when a word changes form depending on the other words to which it relates (e.g., adjectives, adverbs, verbs). In addition, as linguistic criteria (Brazilian Portuguese Language), words that had one of the following characteristics were excluded: (1) Syllable reduction (for example, the word “fósforo” (phosphorus) pronounced as [fósforu]~[fosfru]). (2) Open and close vowels (for example, the word “bolacha” (cookie) pronounced as [ô] and “bola” (ball) pronounced as [ó]). (3) Words with diphthong and hiatus [for example, the word “ vaidade” (vanity) can be pronounced as “vai.da.de,” “va.i.da.de”] and/or monotongation [for example, the word “caixa” (box) can be pronounced as c[a]xa, c[a]xa]. (4) Words with nasal vowels [for example, “orgão” (organ), “homem” (man)]. (5) Tonicity of syllables containing vowel sounds (word selection was made based on the stressed syllable position, and the stressed syllable was in the same position in the target word and in word in the correct answer. (6) Neutralization [e.g., the word “pepino” (cucumber) can be pronounced as “p[e]pino” or “p[i]pino”]. (7) Consonant vocalization (e.g., the pronunciation of words with a consonant corresponding to a post-vowel velar phoneme/l/ may change, and thus it can be pronounced as /u/ or /w/). Most of the words used in Brazilian Portuguese had simple syllable structure, such as consonant-vowel, consonant-vowel-consonant, consonant-consonant-vowel. The screening protocol developed was applied to first graders.

The second goal of this study was to identify the predictive variables of this protocol using Principal Component Analysis. This study was approved by the Research Ethics Committee of the University *Júlio de Mesquita Filho* (FFC/UNESP, São Paulo State University - School of Philosophy and Sciences), Protocol No. 0663/2013.

## Participants

A total of 149 children, aged from 6 years to 6 years and 11 months, of both genders, who were enrolled in the 1st grade of elementary public schools participated in this study. Parents and/or guardians of all the participants signed an informed consent form. Exclusion criteria for participation in the study were as follows: children with sensory, motor, or cognitive impairment and children whose parents/guardians did

not sign the Informed Consent form; inclusion criteria: children whose parents/guardians signed the Informed Consent form and children without sensory, motor, or cognitive impairment, according to information in the school records. Two schools with similar socio-economic status and high rating level in the Secretaria da Educação do Estado de São Paulo (2014) (System of Evaluation of School Performance of the State of São Paulo) participated in this study.

## Procedures

All participants were submitted to the Screening Protocol for Early Identification of Reading Problems (Capellini et al., 2015). The protocol was applied individually in a 50-min session. The protocol was composed of seven cognitive-linguistic tests. Each test was composed of two training trials and test stimuli. The training trials were not scored. During the training trials, the children were informed that the Examiner could offer further explanation about what was being asked and that the Examiner could repeat the stimulus, if necessary. During the test, the Examiner explained that the stimulus could be repeated only once. The rating scale values for Punctuation were: “one” for a correct answer and “zero” for an incorrect answer or a blank. Children marked their answers on an Answer Sheet. The screening protocol was composed of the following tests:

- (1) Letter-naming test. Letters of the alphabet were presented randomly to the children, and they were asked to name the letter shown. Children were presented with visual stimuli with 12 pt Arial uppercase letters. A total of 23 stimuli were shown.
- (2) Test of phonological awareness. Stimuli were orally presented without visual cues. This test was composed of the following subtests:
  - (2.1) Subtest of Rhyme Production. The Examiner presented a word, and the student was asked to say a word that ended with the same sound. This subtest comprised 20 words (target stimuli). Example: target word: “cola” (glue) and expected answer: “bola” (ball).
  - (2.2) Subtest of Rhyme Identification. The Examiner presented a series of three words, and the student was asked to provide a pair of rhyming words (words that sound the same at the end). Twenty groups of three words were presented. Example: a series of three words, “milho” (corn), “baleia” (whale), and “filho” (son). Expected answer: “milho/ filho.”
  - (2.3) Subtest of Syllabic Segmentation. Students were presented with twenty-one words one at a time. The words were selected according to number of syllables (from 2 to 4 syllables), and the children were asked to divide them into syllables. Example: target stimulus: “vaca” (cow). Expected answer: “va – ca.”
  - (2.4) Subtest of Production of words from a given phoneme. A sound/phoneme of the alphabet was presented to the students, and they were asked to say a word beginning with the same sound. Example: target stimulus: phoneme /a/. Expected answer: “asa” (wing).



**TABLE 1 |** List of variables and justification for the development of the screening protocol for early identification of children at risk for dyslexia (Capellini et al., 2017).

Tests of the screening protocol	Justification for test selection
(1) Letter-naming	Pennington and Lefly (2001) stated that Letter-name knowledge is at the intersection between spoken and written language because letters are the written representations of phonemes or combinations of phonemes. It is plausible that the ability to learn letter names depends on underlying phonological development. As pointed out by Share (1995), letter names are, after all, non-words and the ability to repeat and remember non-words.
(2) Phonological Awareness: composed by the subtests (2.1) Rhyme production (2.2) Rhyme identification (2.3) Syllabic segmentation (2.4) Production of words from a given phoneme (2.5) Phonemic synthesis (2.6) Phonemic analysis	As reported by Brunswick et al. (2012), whereas awareness of larger phonological units, such as the syllable and onset-rime, develop independently of reading instruction in 3-to 5-year-olds (Bradley and Bryant, 1983; Badian, 2001; Gipstein et al., 2001), awareness of smaller units of sound, such as the phoneme, usually develop later as a result of reading development (De Jong and van der Leij, 1999; Cardoso-Martins and Pennington, 2004; Ziegler and Goswami, 2006).
(3) Phonological working memory (repetition of words and non-words)	Articulatory loop (phonological working memory system) is thought to be responsible for the temporary storage of verbal information, while other cognitive tasks, such as verbal reasoning or auditory and reading comprehension, are performed (Baddeley, 1986). The task of repeating single non-words is particularly appropriate for use with young children. Throughout the course of childhood, the child, who seems to be innately equipped with both, desire and facility to learn new words, encounters many thousands of unfamiliar words. Exposure to unfamiliar phonological forms is a natural and common occurrence for the child. A further issue of interest is whether non-word repetition ability shares a developmental link with reading achievement. Previously reviewed findings indicate that such a relationship might well exist. Impaired non-word repetition skills have consistently been shown to be characteristic of poor readers and children classified as dyslexic, as described by Gathercole et al. (1994)
(4) Rapid naming speed	According with Jones et al. (2010), naming-speed on these tasks is proposed to assess low-level factors involved in reading fluency, such as: attention to the stimulus bi-hemispheric visual processes responsible for feature detection; matching of feature and pattern encoding to stored orthographic representations; integration of visual and phonological information; and motor activation leading to articulation. Non-alphanumeric stimuli are preferred for use with young students or those who may not have learned letters and digits well enough to be "highly familiar" with them (Wolf and Bowers, 1999; Jones et al., 2010; Kirby et al., 2010).
(5) Silent reading	Measures of silent reading typically include a decision component, such as semantic categorization, sentence verification, or lexical decision. In addition, the main goal in silent reading is to comprehend and assimilate the meaning of the text, which relies on the grapheme-to-semantic decoding in the lexical route (Galin et al., 1992; van den Boer et al., 2014; Zhao et al., 2016). This protocol was based on semantic categorization via lexical processing. Whereas phonological representations might be activated in both silent and oral reading, computation of a phonetic code is specific to oral reading (van den Boer et al., 2014).
(6) Reading words and non-words	Given the nature of the written language, in order to learn to read, the beginning reader needs to decode the written words into speech units and then comprehend the words (and sentences/discourse) to derive meaning, according to The Simple View of Reading (Gough and Tunmer, 1986; Hoover and Gough, 1990). According to this view, reading is composed of word recognition and language comprehension. Word recognition is the translation of print into language (i.e., sounds and words) and comprehension is the making sense of language (Rack et al., 1992; Catts and Hogan, 2003). Moreover, word recognition is a good predictor of reading comprehension performance in children at the beginning of reading acquisition (Chaves-Sousa et al., 2016).
(7) Auditory comprehension of sentences from picture.	Montgomery (1995) argued that sentence comprehension requires that previous information be stored temporarily while new, incoming information is processed. Clark and Clark (1977) have proposed that phonological memory is critical to comprehension because listeners presumably store entire sentences in a phonological input store until all syntactic and semantic analyses have been completed.

(2.5) Subtest of Phonemic Synthesis. Twenty-one words were presented one at time to the students. The Examiner pronounced each phoneme of the word and asked the children to combine them forming a word. The words were selected according to the number of syllables (from 2 to 4 syllables). Example, target stimuli: /k/ /a/ /f/ ε/. Expected answer: "café"/ coffee.

(2.6) Subtest of Phonemic analysis. Twenty-one words were presented one at time to the students. The Examiner pronounced the words and asked the children to divide

them by pronouncing each phoneme/sound of each word. The words were selected according to the number of syllables (from 2 to 4 syllables). Example, target stimulus: "bola/ ball." Expected answer: /b/ /ó/ /l/ /a/.

(2.7) Subtest of Identification of the initial sound/phoneme. Twenty-one words were presented one at time to the students. The Examiner pronounced a word and asked the children to say the initial sound/phoneme of the word out loud. Example, target stimulus: "boca/mouth." Expected answer: /b/.



- (3) Subtest of Phonological Working memory. Twenty-four non-words were presented one at time to the students, and they were asked to repeat the words as they heard them. Repetition was allowed only once. The non-words varied in length (from 1 to 6 syllables).
- (4) Test of Rapid Naming Speed using pictures. Children were presented with flashcards containing seven lines and five columns with colorful picture images: car, ball, duck, house, and key. First, the students were asked to identify each picture stimulus, aiming to verify their recognition. Afterward, they were instructed to say the name of each stimulus as quickly as possible. The examiner used a chronometer to record the time.
- (5) Silent reading. Students were presented with 10 colored pictures, below which there were two words, and they were asked to identify the word that described the picture. The words in the picture belonged to different semantic categories.
- (6) Reading of words and non-words. A list with 20 words and 20 non-words was presented to the students, and they were asked to read them aloud. Children were presented with visual stimuli with 12 pt Arial uppercase letters. A total of 23 stimuli were shown.
- (7) Auditory Comprehension of sentences from pictures. Twenty incomplete sentences with illustrations or figures were presented to the students. The sentences with a missing part were orally presented by the examiner. The children should listen to the sentence and name the missing part, according to the picture.

## RESULTS

Statistical analysis was carried out using the SPSS (Statistical Package for Social Sciences), version 23.0. Some descriptive statistics are shown in **Table 2**.

Principal Components Analysis (PCA) was carried out to reduce the set of Protocol variables before determining the number of variables that could contribute to the early identification of children with dyslexia, such as skills that are predictors of reading acquisition (**Table 3**); Rotation Method used: Varimax with Kaiser Normalization. All factor loadings greater than or equal to 1.00 were used for interpretation.

Although 13 components were retained (factors), only 4 accounted for 64.45% of the total variance (eigenvalues > 1). There was a slight change in all variables due to varimax rotation. Analyzing each factor individually, it was found that the first factor explained 32.56% of variance with no rotation and 23.25% with rotation. The second factor explained 16.06% with no rotation and 15.55% with rotation. The third factor explained 8.05% with no rotation and 14.52% with rotation, and the fourth factor explained 7.77% with no rotation and 11.10% with rotation. In order to clearly define the groups of variables, a correlation matrix was created employing varimax rotation (a more conservative approach) (**Table 4**).

The first factor, called “pre-reading” had high loadings for five variables, indicating 23.26% of variance for all 13 Protocol variables. The variables Letter-naming ( $M = 21.66/SD = 2.88$ ) and Rhyme production ( $M = 11.42/SD = 6.96$ ) had the same loading factor, followed by the variables Rhyme identification, Production of words from a given phoneme, and Identification of the initial phoneme. Letter-naming is one of the most important findings referred to as the foundation of other skills in the first school years, when additional skills are developed. This variable has been proved to be influenced by family environment and pre-school literacy instruction. It can be observed the standard deviation of this variable was low for all students. Moreover, Letter-naming has been associated with phonological awareness because most letter names contain clues regarding their corresponding sound. Rhyme Production ( $M = 11.42/SD = 6.96$ ) and Rhyme identification ( $M = 15.45/SD = 5.09$ ) allow students to realize that words can share identical sound segments, as the perception of greater amounts of sounds will facilitate the formation and increase of lexical and semantic memories, which will be accessed to retrieve auditory information and reading comprehension, afterward. Studies have pointed out that the acquisition of rhymes can occur before literacy instruction, around 3 years old, and it can be combined with skills related to the identification of the initial phonemes, contributing to foster phonemic awareness. Thus, phonemic awareness can emerge as the perception of the smaller segments of spoken words (phonemes), allowing children to perform tasks such as production of words from a given phoneme ( $M = 14.23/SD = 8.25$ ) and identify of the initial phoneme ( $M = 15.45/SD = 5.09$ ), as well to start establishing phoneme-grapheme correspondence, which is important for reading acquisition. The larger standard deviation of these tests suggests that some of the students evaluated may have had difficulties in accessing a word or a phoneme.

The second factor was called “decoding” since it refers to the ability of using grapheme-phoneme correspondences required to read words. It had four variables with 15.56% of variance. The variable Phonemic analysis had the highest factor loadings, followed by the loadings of positive sign of the variables Phonemic Synthesis and Reading words and non-words and the loading of negative sign of the variable Rapid Naming Speed using pictures, which had negative correlation with the second component. In the RAN test, the naming time is measured and the score is expressed in seconds. Lower scores indicate better performance on the test. The ability to identify phonemes contributes to alphabet comprehension since a phoneme may be represented by a sequence of letters. However, it is important to highlight that the Brazilian Teaching Method does not emphasize teaching letter-sound correspondence. Therefore, it can be said that with regard to the Phonemic Analysis ( $M = 2.76/SD = 5.66$ ) and Phonemic Synthesis ( $M = 2.72/SD = 4.84$ ), the students had low performance, suggesting that these variables are important predictors for early identification of dyslexia for this Protocol. These performance difficulties also influenced the students’ performance in the Naming Speed Task ( $M = 43.44/SD = 10.63$ ).

**TABLE 2 |** Distribution of mean (*M*), standard deviation (*SD*), minimum (*min*), and maximum values (*max*), and students' test scores at the 25th and 75th percentiles using the proposed protocol.

Tests	<i>N</i>	Minimum	Maximum	<i>M6</i>	<i>SD</i>	25%	Median	75%
LN (/23)	149	5.00	23.00	21.66	2.88	21.00	23.00	23.00
RP (20)	149	0.00	20.00	11.42	6.96	5.00	13.00	18.00
RI (/20)	149	0.00	20.00	15.45	5.09	12.50	18.00	19.00
SS (/21)	149	0.00	21.00	19.38	3.41	19.00	21.00	21.00
PWPh (/21)	149	0.00	22.00	18.68	3.53	18.00	20.00	21.00
PhS (/21)	149	0.00	21.00	2.76	5.66	0.00	0.00	1.00
PhA (/21)	149	0.00	20.00	2.72	4.84	0.00	0.00	4.00
IPh (/21)	149	0.00	21.00	14.23	8.25	5.00	20.00	21.00
WM (/24)	149	1.00	24.00	20.54	3.05	19.50	21.00	23.00
RAN (seconds)	149	20.00	80.00	43.44	10.63	35.00	42.00	51.00
SR (/10)	149	2.00	10.00	8.95	1.52	8.00	10.00	10.00
RWNW (/40)	149	0.00	40.00	17.15	16.08	0.00	16.00	34.00
AC (/20)	149	0.00	28.00	19.32	2.51	19.00	20.00	20.00

LN, letter-naming; RP, rhyme production; RI, rhyme identification; SS, syllabic segmentation; PWPh, production of words from a given phoneme; PhS, phonemic synthesis; PhA, phonemic analysis; IPh, identification of the initial phoneme; WM, phonological working memory; RAN, rapid naming speed using pictures; SR, silent reading; RWNW, reading of words and non-words; AC, auditory comprehension of sentences from pictures.

**TABLE 3 |** Principal component analysis with extraction sums of squared loadings and rotation sums of squared loadings (Varimax).

Factor	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.233	32.564	32.564	4.233	32.564	32.564	3.024	23.259	23.259
2	2.089	16.069	48.632	2.089	16.069	48.632	2.023	15.565	38.824
3	1.047	8.053	56.685	1.047	8.053	56.685	1.888	14.526	53.350
4	1.010	7.770	64.455	1.010	7.770	64.455	1.444	11.105	64.455
5	0.857	6.592	71.047						
6	0.746	5.736	76.782						
7	0.606	4.665	81.447						
8	0.571	4.394	85.841						
9	0.485	3.734	89.575						
10	0.427	3.288	92.863						
11	0.368	2.827	95.690						
12	0.321	2.472	98.162						
13	0.239	1.838	100.000						

since slow processing speed indicates that the student had difficulties in combining visual with phonological information, suggesting difficulties in reading words and non-words ( $M = 17.15/SD = 16.08$ ).

The third factor, called “Reading” had only two variables accounting for 14.56% of variance. Syllabic segmentation ( $M = 19.38/SD = 3.41$ ) and Silent reading ( $M = 8.95/SD = 1.52$ ) had higher factor loadings. Results indicated that the students had good performance in the Syllabic segmentation test, which does not depend on explicit instruction, and even preschoolers or illiterates have these skills. Silent reading test was the other variable correlated with this factor, which is a specific task that compares recognition between two words and that can be performed by readers without explicit syllable decodification, and the syllables may act as perceptual units in word recognition because of their phonological and orthographic properties, as

mentioned by Ashby (2016). Finally, the fourth factor was called “Auditory processing” and comprised the two last variables, accounting for 11.10% of variance and 64.45% of cumulative variance. Auditory Comprehension of sentences from picture ( $M = 19.32/SD = 2.51$ ) and the Phonological Working memory ( $M = 20.54/SD = 3.05$ ) comprise the last factor. These two variables are somehow correlated because the first one requires that previous information be temporarily stored in the phonological memory (phonological input store until all syntactic and semantic analyses have been completed). **Table 5** shows the distribution of factors, according to the risk criteria – performance below the 25th percentile and at least 51% of the variables correlated with the factors represented by total variance explained by each factor. It can be seen from **Table 5** that 34 students were identified by the first factor, 87 students by the second factor, 19 by the third factor, and 16 by the fourth factor.

**TABLE 4 |** Rotated factor loadings in PCA for the four components.

Variable	Pre-reading	Decoding	Reading	Auditory processing
LN (/23)	<b>0.769*</b>	0.177	0.134	−0.226
RP (20)	<b>0.769*</b>	0.000	0.060	0.254
RI (/20)	<b>0.756*</b>	0.076	0.172	0.239
SS (/21)	0.101	−0.064	<b>0.815*</b>	0.259
PWPh (/21)	<b>0.722*</b>	0.261	0.128	0.002
PhS (/21)	0.268	<b>0.515*</b>	0.014	0.408
PhA (/21)	0.067	<b>0.856*</b>	−0.005	0.140
IPh (/21)	<b>0.687*</b>	−0.123	−0.012	0.444
WM (/24)	0.367	0.192	0.447	<b>0.470*</b>
RAN (seconds)	−0.125	<b>−0.564*</b>	−0.452	0.181
SR (/10)	0.172	0.339	<b>0.703*</b>	−0.177
RWNW (/40)	0.018	<b>0.657*</b>	0.507	−0.100
AC (/20)	0.092	0.027	0.003	<b>0.724*</b>

LN, letter-naming; RP, rhyme production; RI, rhyme identification; SS, syllabic segmentation; PWPh, production of words from a given phoneme; PhS, phonemic synthesis; PhA, phonemic analysis; IPh, identification of the initial phoneme; WM, phonological working memory; RAN, rapid naming speed using pictures; SR, silent reading; RWNW, reading of words and non-words; AC, auditory comprehension of sentences from picture. \* $p < 0.05$ .

## DISCUSSION

This research presented two studies. In study 1, the results indicated that it was possible to develop a screening protocol for early identification of children at risk for dyslexia in first-grade students, using Brazilian Portuguese stimuli. As for Study 2, Principal Component Analysis revealed four factors accounting for 64.45% variance in all Protocol variables. These factors are

consistent with those reported in the National and International literature, and they have been associated with early signs of dyslexia.

Learning how to read in alphabetic systems require the acquisition and domain of associates each distinctive element of visual symbols onto units of sound (phonology). This correspondence process is called phonological recoding (Share, 1995). Nevertheless, this mapping process is influenced by inconsistency in the symbol-to-sound mapping of orthographies. For example, in some Languages it's possible to notice that one letter or letter cluster can be associated with several sound pronunciations (e.g., English, Danish), whereas in other Languages, such as Italian and Spanish, there is a one-to-one correspondence (one letter per sound). However, in some Languages, such as Portuguese and French, it is possible to find both irregularities and regularities, affecting recoding accuracy, which is in line with the reduced consistency of these languages (Ziegler and Goswami, 2006).

The first factor found was denominated “pre-reading” because its variables can be observed before formal education. The “pre-reading” factor comprised the following tests: Letter-naming, Rhyme Production, Rhyme identification, Production of words from a given phoneme, and Identification of the initial phoneme. Letter naming has been considered as a major indicator because its possibility the association between a letter and sound (letter-to-speech sound integration), which can be impaired in children with dyslexia. Although letter naming is considered to be one of the most important predictors of succeeding reading acquisition. However, it's important to note that it is strongly influenced by others factors, such as verbal abilities, teaching methods, and parental input. Letter naming is also closely correlated

**TABLE 5 |** Distribution of the factors according to the risk criteria (performance below the 25th percentile).

Factors	Tests	<i>n</i>	<i>n</i> (risk criteria <25%)	Minimum	Maximum	<i>M</i>	<i>SD</i>
Pre-reading	LN (/23)	149	36	5	23	21.66	2.88
	RP (20)	149	38	0	20	11.42	6.96
	RI (/20)	149	35	0	20	15.45	5.09
	PWPh (/21)	149	44	0	22	18.68	3.53
	IPh (/21)	149	43	0	21	14.23	8.25
	Total		34				
Decoding	PhS (/21)	149	108	0	21	2.76	5.66
	PhA (/21)	149	84	0	20	2.72	4.84
	RAN (seconds)	149	145	20	80	43.44	10.63
	RWNW (/40)	149	75	0	40	17.15	16.08
	Total		87				
Reading	SR (/10)	149	39	2	10	8.95	1.52
	SS (/21)	149	43	0	21	19.38	3.41
	Total		19				
Auditory processing	WM (/24)	149	36	1	24	20.54	3.05
	AC (/20)	149	35	0	28	19.32	2.51
	Total		16				

LN, letter-naming; RP, rhyme production; RI, rhyme identification; SS, syllabic segmentation; PWPh, production of words from a given phoneme; PhS, phonemic synthesis; PhA, phonemic analysis; IPh, identification of the initial phoneme; WM, phonological working memory; RAN, rapid naming speed using pictures; SR, silent reading; RWNW, reading of words and non-words; AC, auditory comprehension of sentences from picture.

with phonological awareness (Lerner and Lonigan, 2016). The performance of letter knowledge and phonological awareness at kindergarten have been strongly referred as predictors for First-grade reading achievement. These findings were pointed even when variables, such as parental education level and teacher-rated academic competence (Ortiz et al., 2012; Lerner and Lonigan, 2016); Lerner and Lonigan (2016) also discussed the influence of phonological awareness on the acquisition of letter knowledge.

Unfortunately, even though international researchers have pointed out the role of Letter-naming and teaching of letter-naming correspondence in several alphabetic languages, according to the *Parâmetros curriculares nacionais da Língua Portuguesa* (Brasil, 1997) (National curricular parameters of Portuguese Language), the current understanding of the relationship between writing acquisition and writing skills confront entrenched beliefs that the phonics instruction domain is a prerequisite for language teaching, indicating that the two learning processes (literacy and language teaching itself) could occur simultaneously. Therefore, with regard to the Alphabetic language principles, the acquisition of alphabetic knowledge does not guarantee that the student will be able to understand or produce texts in writing. Finally, according to these parameters, teaching basic units could comprise not only reading comprehension, which does not mean that teaching words or sentences would not focus on specific didactic situations that would benefit students.

Perhaps, because of this lack of systematic approach to teaching, letter-naming skills have still been considered as one of the predictors to dyslexia, as reported in international studies on alphabetic language. Kim et al. (2010) and Lerner and Lonigan (2016) argued that for letter names that incorporated important traces of the corresponding sound. Therefore, knowing the name of a letter was a strong predictor to realize if the children knows the corresponding sound. However, this can be observed in children with good developed phonological awareness skills.

According to Lonigan et al. (2009), phonological awareness develops along a *continuum* awareness of large and concrete sound units (i.e., words, syllables) to awareness of small and abstract sound units (i.e., phonemes). The other variables correlated with Factor 1 concern the perception of large amounts of sounds (Rhymes) and the use of phoneme knowledge. The findings of this study showed that students had more difficulties with Identification of the Initial Phoneme than with Production of words from a given phoneme. Therefore, in these phonemic tests, students have to identify the first phoneme of the words (i.e., alliteration) and retrieve another word from the phonological long-term memory. Our results corroborate with those found in the literature suggesting that even before entering preschool, children learn some basic language skills and notions (detection, rhyming, and alliteration) that will facilitate the development of reading skills based on a variety of life experiences. These experiences contribute to their acquisition of receptive vocabulary phonological skills, and narrative understanding and production (Hayes and Slater, 2008; Manz et al., 2010). With regard to Phonemic awareness, as mentioned by Silvén et al. (2002), this finding may support the assumption that conscious access to speech patterns is

influenced, at least indirectly, by advances in implicit phonetic and phonotactic representations that can be related to language development during the 1st year of life. Ouellette and Haley (2013) stated that the principal motivation for considering the role of vocabulary in the emergence of phonemic awareness could be associated with the first words stored in mental lexicon. As new words are added, segmental representation becomes necessary so that similar sounding items are not confused with each other. Essentially, increased extensiveness of oral vocabulary causes restructuring, by which there are more specific phonemic-level representations. Accordingly, Law et al. (2016) evaluated a group of pre-reading children with a family risk for dyslexia. As results, the authors founded that there was an influence of phonological and morphological awareness on reading development. According to Morris et al. (2012), morphological awareness can be defined as the explicit awareness and ability to manipulate and reflect upon the morphemic structure of words, which has already been demonstrated in pre-reading children. The results obtained suggest that phonological awareness is a relevant component of morphological awareness, independent of reading experience. It is also important to highlight that in the present study, the variables correlated with factor 2, corroborate those reported in the study of Hulme et al. (2015), who found that children at risk for dyslexia show general deficits in oral language skills in the preschool years. Those deficits are presented in a way that a percentage of these children satisfy the criteria for language impairment diagnosis. Poor oral language skills, in turn, appear to affect the later development of decoding (through problems in acquiring letter-sound knowledge and phoneme awareness) as well as reading comprehension abilities. Based on these international studies, it can be said that the variables correlated with the first factor proved important as predictive variables in the Brazilian Portuguese Language. As an alphabetic language, difficulties in acquiring letter naming and initial phonological awareness skills can be seen as a sign of reading difficulties.

Thus, the 2nd factor was comprised the following tests: Phonemic Synthesis, Phonemic analysis, Rapid Naming Speed using pictures, and Reading of words and non-words. As described by Ouellette and Haley (2013), phonemic awareness can also be categorized based on how it is being used. Specifically, explicit awareness at the level of the phoneme includes both analytic (ability to break a word down into constituent sounds) and synthetic skills (combining sounds together to make a larger segment, such as word). Analysis tasks are more difficult than synthesis tasks. Our findings showed that phonemic analysis and phonemic synthesis had different loadings; however, the students had similar performance (mean) on the tests of phonemic analysis and phonemic synthesis. Furthermore, our findings also showed difficulties in reading words and non-words and a negative loading for Rapid Naming Speed (RAN). This suggests that difficulties in decoding skills were related with slow phonological access. Evaluating the double deficit hypothesis, Brazilian and International studies (Wolf and Bowers, 1999; Andrade et al., 2013; Silva and Capellini, 2013; De Groot et al., 2017), demonstrated the relationship among phonological awareness difficulties, dyslexia, and impaired RAN. Hulme et al.



(2015) argued that phoneme awareness and letter knowledge are the most important predictors of early word-reading skills across several languages, and there is evidence of reciprocal interaction between them. Extending these ideas mentioned before (Shaywitz and Shaywitz, 2005; Kirby et al., 2010; Snowling, 2013; Hulme et al., 2015). Kirby et al. (2010) examined RAN effects across languages and the impact of its relationship to reading. These authors also reviewed the instructional literature aiming to improve and to use RAN as a predictor of RTI. They concluded that RAN is uniquely associated with a variety of reading tasks across orthographies, and that the use of RAN measures would be very useful for early identification.

The third factor was called “reading” and had only two variables. It was composed of the tests Syllabic segmentation and Silent reading. Syllabic segmentation is one of the skills that does not depend on explicit instruction. Our results might suggest that students had good performance on Syllabic segmentation, but they had some difficulties in the reading tests. One possible explanation is that in the test of silent reading they could use the visual or orthographic routes instead of the phonological route. Thus, the orthographic process occurs when groups of letters or entire words are processed as single units rather than as a sequence of grapheme–phoneme correspondences, which is related to phonological processing (Ehri, 1997; Pinheiro et al., 2008; Oliveira and Capellini, 2010; Mota et al., 2012; Majerus and Cowan, 2016). Moreover, Kirby et al. (2010) stated that because of the orthographic process, it’s possible to establish the mechanism of quick recognition of very frequent or familiar (Morris et al., 1998; Vaessen et al., 2009). Since the reading test was composed of reading words and non-words, both variables correlated with factor 3 may demonstrate that simpler phonological awareness skills (e.g., Syllabic segmentation) can contribute to early identification of dyslexia because syllabic segmentation does not depend on reading instruction and is not related to oral language acquisition. The other variable, the variable related to Reading, will also contribute for early identification of dyslexia since it enables evaluation of the reading level, considering the use of the lexical route and the phonological route. In a study addressing cross-language reading comparison, Ziegler and Goswami (2006) reported that one of the most significant findings was that the students who were acquiring reading in orthographically consistent languages (Greek, Finnish, German, Italian, and Spanish) were close to ceiling in both word and non-word reading by the middle of first grade. Unfortunately, this was not observed in the Brazilian population studied. In contrast, the standard deviation of this test for Brazilian Portuguese Language was large (students who were not able to read a single word). According to Scliar-Cabral (2003), Brazilian Portuguese is also considered to be transparent, and reading can be performed only when the grapheme–phoneme matching rules are known and phonological decoding is used, mostly at the beginning of literacy. However, it is worth highlighting that despite the existence of the National curricular parameters of Portuguese language (Brasil, 1997) there is no systematic teaching of grapheme–phoneme correspondence rules. The findings shown in **Table 5** can be justified by the lack of systematic

teaching, and because of this, factor 2 (“Decoding”) identified a larger number of students, when compared with factors 1, 3, and 4. Factor 2 was composed of the Phonemic tests Rapid Naming Speed using pictures and Reading of words and non-words.

The 4th factor comprised the tests of Auditory Comprehension of sentences from pictures and Phonological Working memory. A deficit in verbal short-term memory is well documented in dyslexia and can be observed in tasks such as reading longer pseudowords or sequences are used, specially repeating 4 to 6-syllable pseudowords that are related with phonological deficits (Ramus and Ahissar, 2012). Baddeley (1986) defined phonological memory as the coding of information, a sound-based representation system for temporary storage, that can be measured by immediate recall of verbally presented material (e.g., repetition of non-words), as also reported by Lonigan et al. (2009). Studies (Ehri, 1997; Pinheiro et al., 2008; Oliveira and Capellini, 2010; Mota et al., 2012; Majerus and Cowan, 2016) reported that dyslexia children have difficulties in phonological awareness, and this difficulties can still be observed in adults with a history of dyslexia. This difficulty plays an important role to characterize the dyslexia profile, suggesting that the reduction of the amount of phonological and graphemic information that can be co-activated during the reading process can influence the recoding reading process, when grapheme–phoneme correspondence are not yet automatized, leading to difficulties in reading comprehension. Therefore, it’s important that the children have an efficient phonological memory that enable the maintenance of an accurate representation of the correspondence grapheme–phonemes while word decoding and, consequently, allocate more cognitive resources to comprehension processes. (Lonigan et al., 2009). Spoken-language comprehension and processing depend on the accurate isolation and interpretation of meaningful units of speech such as words, sentences, or utterances. Such high-level perceptual units correspond to the consolidation of basal acoustic-phonetic cues that can be categorized within various time scales corresponding to various phonological grain size units. Therefore, in order to have a good performance in Auditory Comprehension of sentences from pictures, the students evaluated were able to decode phonological information.

Finally, it is possible to identify the variables for the Screening Protocol for Early Identification of Reading Problems in Brazilian children enrolled in first grade, according to the order of the predictive value of each variable, which was as follows: Letter-naming, Rhyme Production, Rhyme identification, Production of words from a given phoneme, Identification of the initial phoneme, Phonemic analysis, Reading of words and non-words, Phonemic Synthesis, Rapid Naming Speed using pictures, Syllabic segmentation, Silent reading, Auditory Comprehension of sentences from picture, and Phonological Working memory. Combined, these three factors (Pre-reading, Decoding, and Reading) accounted for 53.35% of the students’ performance on the Protocol, and thus these factors would be statically sufficient to create a version of the protocol proposed. However, future

studies are necessary to verify the exclusion of the 4th factor (Auditory processing).

Our findings are in agreement with those found by Thompson et al. (2015), who indicated that early identification of “reading problems” is difficult, and the development of assessment protocols for this age and grade level are extremely important, as they can prevent future learning damages. Furthermore, our findings also suggest that early language problems can be considered as risk factors for dyslexia, but they can be also considered as risk factors for this disability for children entering school. Although Principal components Analysis revealed four factors, it is important to highlight that future analysis are still necessary to investigate the underlying factors affecting test items. However, if we take into account the educational reality in Brazil, the screening protocol proposed accomplished one of its main goals, which is helping professionals, such as teachers, Speech Language Therapists, and others to identify students at risk for dyslexia or other reading problems in 1st grade, since this type of protocols are practically non-existent in the country. In Brazil, one of the issues related to the identification of children at risk for dyslexia is the long period of time until they are referred to diagnostic centers. This results is consistent with the view that many children who have language delay and receive proper treatment can learn how to read. However, it is worth mentioning that these children continue at risk of having difficulties in reading skills and can present others difficulties, including social problems.

## CONCLUSION

Results indicated that screening protocol developed in the present study showed four major factors: pre-reading (Letter-naming, Rhyme production, Rhyme identification, Production of words from a given phoneme, Identification of the initial phoneme); decodification (Phonemic synthesis, Phonemic analysis, Rapid Naming Speed using pictures, Reading of words and non-words);

reading (Silent Reading and Syllabic segmentation); and Auditory processing (Phonological working memory and comprehension of sentences from pictures) to identify Brazilian Portuguese speaking children at risk for dyslexia.

Based on the PCA carried out, our findings showed the effective use of the proposed Screening Protocol to analyze the predictive factors that can explain later reading achievement.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of ‘Frontiers guidelines’ 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 ‘Ethics Committee of Faculty of Philosophy and Sciences, São Paulo State University “Júlio de Mesquita Filho” (FFC/UNESP).’

## AUTHOR CONTRIBUTIONS

GG, AC, and SC had substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; drafting the work or revising it critically for important intellectual content; final approval of the version to be published; agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## ACKNOWLEDGMENTS

The authors are grateful for the financial support provided by CNPq, The National Council for Scientific and Technological Development (Universal Notice MCT/CNPq number 14/2012).

## REFERENCES

- Alloway, T. P., Gathercole, S. E., Adams, A.-M., Willis, C., Eaglen, R., and Lamont, E. (2005). Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. *Br. J. Dev. Psychol.* 23, 417–426. doi: 10.1348/026151005X26804
- Andrade, O. V., Andrade, P. E., and Capellini, S. A., (2013). Identificação precoce do risco para transtornos da atenção e da leitura em sala de aula. *Psicol. Teoria Pesqui.* 29, 167–176. doi: 10.1590/S0102-37722013000200006
- Andrade, O. V. C. A., Prado, P. S. T., and Capellini, S. A., (2011). Desenvolvimento de ferramentas pedagógicas para identificação de escolares de risco para a dislexia. *Rev. Psicopedag.* 28, 14–28.
- Araújo, S., Faísca, L., Reis, A., Marques, J. F., and Petersson, K. M. (2016). Visual naming deficits in dyslexia: an ERP investigation of different processing domains. *Neuropsychologia* 91, 61–76. doi: 10.1016/j.neuropsychologia.2016.07.007
- Araújo, S., Pacheco, A., Faísca, L., Petersson, K. M., and Reis, A. (2010). Visual rapid naming and phonological abilities: different subtypes in dyslexic children. *Int. J. Psychol.* 45, 443–452. doi: 10.1080/00207594.2010.499949
- Ashby, J. (2016). “Why does prosody accompany fluency? Re-conceptualizing the role of phonology in reading,” in *Reading Fluency*, eds A. Khateb and I. Bar-Kochva (Berlin: Springer International Publishing), 65–89. doi: 10.1007/978-3-319-30478-6\_5
- Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.
- Badian, N. A. (2001). Phonological and orthographic processing: their roles in reading prediction. *Ann. Dyslexia* 51, 179–202. doi: 10.1007/s11881-001-0010-5
- Bradley, L., and Bryant, P. E. (1983). Categorizing sounds and learning to read: a causal connection. *Nature* 301, 419–421. doi: 10.1038/301419a0
- Brasil (1997). *Secretaria de Educação Fundamental. Parâmetros Curriculares Nacionais: Língua Portuguesa*. Brasília: Secretaria de Educação Fundamental.
- Brizzolara, D., Chilosi, A., Cipriani, P., Di Filippo, G., Gasperini, F., Mazzotti, S., et al. (2006). Do phonologic and rapid automatized naming deficits differentially affect dyslexic children with and without a history of language delay? A study of Italian dyslexic children. *Cogn. Behav. Neurol.* 19, 141–149. doi: 10.1097/01.wnn.0000213902.59827.19
- Brunswick, N., Martin, G. N., and Rippon, G. (2012). Early cognitive profiles of emergent readers: a longitudinal study. *J. Exp. Child Psychol.* 111, 268–285. doi: 10.1016/j.jecp.2011.08.001
- Câmara, J. R. (1970a). *Estrutura da Língua Portuguesa*. Petrópolis: Vozes.



- Câmera, J. R. (1970b). *Problemas de Linguística Descritiva*, 3. Edn. Petrópolis: Vozes.
- Capellini, S. A., César, A. B. P., and Germano, G. D. (2015). Early identification of reading problems: preliminary study with students of 1st grade. *Proc. Soc. Behav. Sci.* 174, 1351–1355. doi: 10.1016/j.sbspro.2015.01.758
- Capellini, S. A., César, A. B. P., and Germano, G. D. (2017). *Protocolo de Identificação Precoce dos Problemas de Leitura*. Booktoy: Livraria e Editora.
- Capellini, S. A., and Conrado, T. L. B. C. (2009). Desempenho de escolares com e sem dificuldades de aprendizagem de ensino particular em habilidade fonológica, nomeação rápida, leitura e escrita. *Rev. CEFAC* 11, 183–193. doi: 10.1590/S1516-18462009005000002
- Capellini, S. A., Sampaio, M. N., Fukuda, M. T. M., Oliveira, A. M., Fadini, C. C., and Martins, M. A. (2009). Protocolo de identificação precoce dos problemas de leitura: estudo preliminar com escolares de 1o ano escolar. *Rev. Psicopedag.* 26, 367–375.
- Cardoso-Martins, C., and Pennington, B. F. (2004). The relationship between phoneme awareness and rapid serial naming skills and literacy acquisition: the role of developmental period and reading ability. *Sci. Stud. Read.* 8, 27–52. doi: 10.1207/s1532799xssr0801\_3
- Carroll, J. M., Mundy, I. R., and Cunningham, A. J. (2014). The roles of family history of dyslexia, language, speech production and phonological processing in predicting literacy progress. *Dev. Sci.* 17, 727–742. doi: 10.1207/s1532799xssr0801\_3
- Catts, H. W., and Hogan, T. P. (2003). Language basis of reading disabilities and implications for early identification and remediation. *Read Psychol.* 24, 223–246. doi: 10.1080/0270210390227314
- Cervera-Mérida, J. F., and Ygual-fernández, A. (2003). Intervención logopédica en los trastornos fonológicos desde el paradigma psicolingüístico del procesamiento del habla. *Rev. Neurol.* 36, 39–53.
- Chard, D. J., and Dickson, S. V. (1999). Phonological awareness instructional and assessment guidelines. *Interv. Sch. Clin.* 34, 261–270. doi: 10.1007/BF02648171
- Chaves-Sousa, S., Santos, S., Viana, F. L., Vale, A. P., Cadime, I., Prieto, G., et al. (2016). Development of a word reading test: identifying students at-risk for reading problems. *Learn Individ. Diff.* 56, 159–166. doi: 10.1016/j.lindif.2016.11.008
- Clark, H., and Clark, E. (1977). *Psychology and Language*. New York, NY: Harcourt Brace Jovanovic.
- 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. doi: 10.1037/0033-295X.108.1.204
- De Groot, B. J., Van den Bos, K. P., Van der Meulen, B. F., and Minnaert, A. E. (2017). Rapid naming and phonemic awareness in children with or without reading disabilities and/or ADHD. *J. Learn. Disabil.* 50, 168–179. doi: 10.1177/0022219415609186
- De Jong, P. F., and Olson, R. K. (2004). Early predictors of letter knowledge. *J. Exp. Child Psychol.* 88, 254–273. doi: 10.1016/j.jecp.2004.03.007
- De Jong, P. F., and Van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: results from a Dutch latent variable longitudinal study. *J. Educ. Psychol.* 91, 450–476. doi: 10.1037/0022-0663.91.3.450
- Denckla, M. B., and Rudel, R. (1976). Rapid “automatized” naming (RAN): dyslexia differentiated from other learning disabilities. *Neuropsychologia* 14, 471–479. doi: 10.1016/0028-3932(76)90075-0
- Duncan, L. G., Castro, S. L., Defior, S., Seymour, P. H., Baillie, S., Leybaert, J., et al. (2013). Phonological development in relation to native language and literacy: variations on a theme in six alphabetic orthographies. *Cognition* 127, 398–419. doi: 10.1016/0028-3932(76)90075-0
- Ehri, L. C. (1997). “Learning to read and learning to spell are one and the same, almost,” in *Learning to spell: Research, Theory, and Practice Across Languages*, eds C. A. Perfetti, L. Rieben, and M. Fayol (Mahwah, NJ: Lawrence Erlbaum Associates), 237–268.
- Fadini, C. C., and Capellini, S. A. (2011). Treinamento de habilidades fonológicas em escolares de risco para dislexia. *Rev. Psicopedag.* 28, 3–13.
- Franceschini, S., Gori, S., Ruffino, M., Pedrollo, K., and Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Curr. Biol.* 22, 814–819. doi: 10.1016/j.cub.2012.03.013
- Fukuda, M. T. M., and Capellini, S. A. (2011). Treinamento de habilidades fonológicas e correspondência grafema-fonema em crianças de risco para dislexia. *Rev. CEFAC* 13, 227–235. doi: 10.1590/S1516-18462010005000074
- Fukuda, M. T. M., and Capellini, S. A. (2012). Programa de Intervenção Fonológica associado à correspondência grafema-fonema em escolares de risco para a dislexia. *Psicol. Refl. Crít.* 25, 783–790. doi: 10.1590/S0102-79722012000400018
- Galin, D., Raz, J., Fein, G., Johnstone, J., Herron, J., and Yingling, C. (1992). EEG spectra in dyslexic and normal readers during oral and silent reading. *Electroencephalogr. Clin. Neurophysiol.* 82, 87–101. doi: 10.1016/0013-4694(92)90151-7
- Gathercole, S. E., and Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: a longitudinal study. *J. Mem. Lang.* 28, 200–213. doi: 10.1016/0749-596X(89)90044-2
- Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., and Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Appl. Cogn. Psychol.* 13, 65–77. doi: 10.1002/(SICI)1099-0720(199902)13:1<65::AID-ACP548>3.0.CO;2-O
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., and Emslie, H. (1994). The children’s test of nonword repetition: a test of phonological working memory. *Memory* 2, 103–127. doi: 10.1080/09658219408258940
- Germano, G. D. (2011). *Instrumento de Avaliação Metafonológica Para Caracterização de Escolares com Dislexia, Transtorno de Dificuldades de Aprendizagem*. Doctorate’s thesis, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília.
- Germano, G. D., and Capellini, S. A. (2008). Eficácia do programa de remediação auditivo-visual computadorizado em escolares com dislexia. *Pró Fono Rev. Atual. Cient.* 20, 237–242. doi: 10.1590/S0104-56872008000400006
- Germano, G. D., and Capellini, S. A. (2011). Performance of students with dyslexia, learning disabilities and learning difficulties in metaphonological abilities tests (PROHFON). *J. Soc. Bras. Fonoaudiol.* 23, 135–141. doi: 10.1590/S2179-64912011000200010
- Germano, G. D., Reilhac, C., Capellini, S. A., and Valdois, S. (2014). The phonological and visual basis of developmental dyslexia in Brazilian Portuguese reading children. *Front. Psychol.* 5:1169. doi: 10.3389/fpsyg.2014.01169
- Gipstein, M., Brady, S., and Fowler, A. (2001). “Questioning the role of syllables and rimes in early phonological awareness,” in *Early Identification and Remediation of Reading Disability*, Second Edn, ed. M. Masland (Parkton, MD: York Press), 179–216.
- Godoy, D. M., and Pinheiro, A. (2013). *O Que Sabemos Sobre A Contribuição da Consciência Fonêmica Para A Aprendizagem Inicial da Leitura e da Escrita. A Aprendizagem da Leitura e da Escrita*. São Paulo: Vetor.
- Gombert, J. E. (2003). “Atividades metalinguísticas e aprendizagem da leitura,” in *Metalinguagem e Aquisição da Escrita: Contribuições da Pesquisa Para a Prática da Alfabetização*, ed. M. R. Maluf (São Paulo: Casa do Psicólogo), 19–63.
- Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104
- Hayes, R., and Slater, A. (2008). Three month-olds’ detection of alliteration in syllables. *Infant Behav. Dev.* 31, 153–156. doi: 10.1016/j.infbeh.2007.07.009
- Hoover, W. A., and Gough, P. B. (1990). *Read Writ*, Vol. 2. Dordrecht: Kluwer Academic Publishers, 127. doi: 10.1007/BF00401799
- Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., and Snowling, M. J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychol. Sci.* 26, 1877–1886. doi: 10.1177/0956797615603702
- Hulme, C., and Snowling, M. J. (2014). The interface between spoken and written language: developmental disorders. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20120395. doi: 10.1098/rstb.2012.0395
- Johnston, T. C., and Kirby, J. R. (2006). The contribution of naming speed to the simple view of reading. *Read. Writ.* 19, 339–361. doi: 10.1007/s11145-005-4644-2
- Jones, M. W., Branigan, H. P., Hatzidaki, A., and Obregón, M. (2010). Is the ‘naming’ deficit in dyslexia a misnomer? *Cognition* 116, 56–70. doi: 10.1016/j.cognition.2010.03.015
- Kim, Y. S., Petscher, Y., Schatschneider, C., and Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *J. Educ. Psychol.* 102, 652–667. doi: 10.1037/a0019643

- 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
- Law, J. M., Wouters, J., and Ghesquière, P. (2016). The influences and outcomes of phonological awareness: a study of MA, PA and auditory processing in pre-readers with a family risk of dyslexia. *Dev. Sci.* 20:12453. doi: 10.1111/desc.12453
- Lerner, M. D., and Lonigan, C. J. (2016). Bidirectional relations between phonological awareness and letter knowledge in preschool revisited: a growth curve analysis of the relation between two code-related skills. *J. Exp. Child Psychol.* 144, 166–183. doi: 10.1016/j.jecp.2015.09.023
- Lonigan, C. J., Anthony, J. L., Phillips, B. M., Purpura, D. J., Wilson, S. B., and McQueen, J. D. (2009). The nature of preschool phonological processing abilities and their relations to vocabulary, general cognitive abilities, and print knowledge. *J. Educ. Psychol.* 101, 345–358. doi: 10.1037/a0013837
- Majerus, S., and Cowan, N. (2016). The nature of verbal short-term impairment in dyslexia: the importance of serial order. *Front. Psychol.* 7:1522. doi: 10.3389/fpsyg.2016.01522
- Manz, P. H., Hughes, C., Barnabas, E., Bracaliello, C., and Ginsburg-Block, M. (2010). A descriptive review and meta-analysis of family-based emergent literacy interventions: to what extent is the research applicable to low-income, ethnic-minority or linguistically-diverse young children? *Early Child. Res. Q.* 25, 409–431. doi: 10.1016/j.jecresq.2010.03.002
- Montgomery, J. W. (1995). Sentence comprehension in children with specific language impairment: the role of phonological working memory. *J. Speech Lang. Hear.* 38, 187–199. doi: 10.1044/jslr.3801.187
- Morris, R. D., Lovett, M. W., Wolf, M., Sevcik, R. A., Steinbach, K. A., Frijters, J. C., et al. (2012). Multiple-component remediation for developmental reading disabilities: IQ, socioeconomic status, and race as factors in remedial outcome. *J. Learn. Disabil.* 45, 99–127. doi: 10.1177/0022219409355472
- Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., et al. (1998). Subtypes of reading disability: variability around a phonological core. *J. Educ. Psychol.* 90, 347–373. doi: 10.1037/0022-0663.90.3.347
- Mota, M. M. P. E., Vieira, M. T., Bastos, R. R., Dias, J., Paiva, N., Mansur-Lisboa, S., et al. (2012). Contextual reading and metalinguistic processing in Brazilian Portuguese: a longitudinal study. *Psicol. Refl. Crít.* 25, 114–120. doi: 10.1590/S0102-79722012000100014
- Norton, E. S., and Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annu. Rev. Psychol.* 63, 427–452. doi: 10.1146/annurev-psych-120710-100431
- Oliveira, A. M. D., and Capellini, S. A. (2010). Desempenho de escolares na adaptação brasileira da avaliação dos processos de leitura. *Pró Fono R. Atual. Cient.* 22, 555–560. doi: 10.1590/S0104-56872010000400033
- Ortiz, M., Folsom, J. S., Otaiba, S. A., Greulich, L., Thomas-Tate, S., and Connor, C. M. (2012). The componential model of reading: predicting first grade reading performance of culturally diverse students from ecological, psychological, and cognitive factors assessed at kindergarten entry. *J. Learn. Disabil.* 45, 406–417. doi: 10.1177/0022219411431242
- Ouellette, G. P., and Haley, A. (2013). One complicated extended family: the influence of alphabetic knowledge and vocabulary on phonemic awareness. *rJ. Read. Res.* 36, 29–41. doi: 10.1111/j.1467-9817.2010.01486.x
- 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
- Peterson, R. L., and Pennington, B. F. (2012). Developmental dyslexia. *Lancet* 379, 1997–2007. doi: 10.1016/S0140-6736(12)60198-6
- Pinheiro, A. M. V., Cunha, C. R., and Lúcio, P. S. (2008). Tarefa de leitura de palavras em voz alta: uma proposta de análise dos erros. *Rev. Port. Educ.* 21, 115–138.
- Rack, J. P., Snowling, M. J., and Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: a review. *Read. Res. Q.* 27, 29–53. doi: 10.2307/747832
- Ramus, F., and Ahissar, M. (2012). Developmental dyslexia: the difficulties of interpreting poor performance, and the importance of normal performance. *J. Cogn. Neurosci.* 29, 104–122. doi: 10.1080/02643294.2012.677420
- Scliar-Cabral, L. (2003). *Princípios do Sistema Alfabético do Português do Brasil*. São Paulo: Contexto.
- Secretaria da Educação do Estado de São Paulo (2014). *Relatório Pedagógico de Língua Portuguesa (Saresp)*. São Paulo: Secretaria da Educação do Estado de São Paulo.
- Selkirk, E. O. (1982). “The syllable,” in *The Structure of Phonological Representations*, eds H. Van der Hulst and N. Smith (Dordrecht: Foris Publications).
- Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition* 55, 151–218. doi: 10.1016/0010-0277(94)00645-2
- Shaywitz, S. E., and Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biol. Psychiatry* 57, 1301–1309. doi: 10.1016/j.biopsych.2005.01.043
- Silva, C. D., and Capellini, S. A. (2013). Desempenho de escolares com e sem transtorno de aprendizagem em leitura, escrita, consciência fonológica, velocidade de processamento e memória de trabalho fonológica. *Rev. Psicopedag.* 30, 3–11.
- Silvén, M., Niemi, P., and Voeten, M. J. (2002). Do maternal interaction and early language predict phonological awareness in 3-to 4-year-olds? *Cogn. Dev.* 17, 1133–1155. doi: 10.1016/S0885-2014(02)00093-X
- Skeide, M. A., Kirsten, H., Kraft, I., Schaadt, G., Müller, B., Neef, N., et al. (2015). Genetic dyslexia risk variant is related to neural connectivity patterns underlying phonological awareness in children. *Neuroimage* 118, 414–421. doi: 10.1016/j.neuroimage.2015.06.024
- Snowling, M. J. (2013). Early identification and interventions for dyslexia: a contemporary view. *JORSEN* 13, 7–14. doi: 10.1111/j.1471-3802.2012.01262.x
- Thompson, P. A., Hulme, C., Nash, H. M., Gooch, D., Hayiou-Thomas, E., and Snowling, M. J. (2015). Developmental dyslexia: predicting individual risk. *J. Exp. Child Psychol.* 56, 976–987. doi: 10.1111/jcpp.12412
- 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 den Boer, M., van Bergen, E., and de Jong, P. F. (2014). Underlying skills of oral and silent reading. *J. Exp. Child Psychol.* 128, 138–151. doi: 10.1016/j.jecp.2014.07.008
- 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
- Wolf, M., and Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexia. *J. Educ. Psychol.* 91, 415–438. 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
- Zhao, J., Kwok, R. K., Liu, M., Liu, H., and Huang, C. (2016). Underlying skills of oral and silent reading fluency in Chinese: perspective of visual rapid processing. *Front. Psychol.* 7:2082. doi: 10.3389/fpsyg.2016.02082
- Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F. X., and Perry, C. (2008). Developmental dyslexia and the dual route model of reading: simulating individual differences and subtypes. *Cognition* 107, 151–178. doi: 10.1016/j.cognition.2007.09.004
- Ziegler, J. C., and Goswami, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Dev. Sci.* 9, 429–436. doi: 10.1111/j.1467-7687.2006.00509.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Germano, César and Capellini. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Effects of a Syllable-Based Reading Intervention in Poor-Reading Fourth Graders

Bettina Müller<sup>1\*</sup>, Tobias Richter<sup>1</sup>, Panagiotis Karageorgos<sup>2</sup>, Sabine Krawietz<sup>3</sup> and Marco Ennemoser<sup>4</sup>

<sup>1</sup> Department of Psychology IV, Educational Psychology, University of Würzburg, Würzburg, Germany, <sup>2</sup> Department of Psychology, University of Kassel, Kassel, Germany, <sup>3</sup> Department of Sport Science, Technical University of Darmstadt, Darmstadt, Germany, <sup>4</sup> Department of Special Education, Ludwigsburg University of Education, Ludwigsburg, Germany

In transparent orthographies, persistent reading fluency difficulties are a major cause of poor reading skills in primary school. The purpose of the present study was to investigate effects of a syllable-based reading intervention on word reading fluency and reading comprehension among German-speaking poor readers in Grade 4. The 16-session intervention was based on analyzing the syllabic structure of words to strengthen the mental representations of syllables and words that consist of these syllables. The training materials were designed using the 500 most frequent syllables typically read by fourth graders. The 75 poor readers were randomly allocated to the treatment or the control group. Results indicate a significant and strong effect on the fluency of recognizing single words, whereas text-level reading comprehension was not significantly improved by the training. The specific treatment effect provides evidence that a short syllable-based approach works even in older poor readers at the end of primary school.

**Keywords:** older poor readers, primary school, word reading fluency, reading comprehension, syllable-based intervention

## OPEN ACCESS

### Edited by:

Simone Aparecida Capellini,  
São Paulo State University, Brazil

### Reviewed by:

Angela Jocelyn Fawcett,  
Swansea University, United Kingdom  
Manuel Soriano-Ferrer,  
Universitat de València, Spain

### \*Correspondence:

Bettina Müller  
bettina.mueller@uni-wuerzburg.de

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 30 January 2017

**Accepted:** 06 September 2017

**Published:** 20 September 2017

### Citation:

Müller B, Richter T, Karageorgos P,  
Krawietz S and Ennemoser M (2017)  
Effects of a Syllable-Based Reading  
Intervention in Poor-Reading Fourth  
Graders. *Front. Psychol.* 8:1635.  
doi: 10.3389/fpsyg.2017.01635

## INTRODUCTION

Recognizing written words and understanding written texts are among the most important competences children need to acquire in primary school. However, according to the Progress in International Literacy Study (PIRLS), 15.4% of fourth graders learning to read in German fail to master fundamental reading comprehension tasks (Bos et al., 2012). These students experience problems constructing a coherent representation of texts, because they are unable to connect information from different parts of the text or draw knowledge-based inferences that are crucial for comprehension. Importantly, most students with reading comprehension problems continue to struggle with basic reading processes at the word-level (Torppa et al., 2007), which leads to difficulties in understanding written texts at a level that is functional to meet the requirements of school and society. Individual differences in reading skills are known to be very stable across the school years. A longitudinal study with German children found that 70% of the children who fall behind in word reading fluency in Grade 1 still score low in reading fluency in Grade 8 (Landerl and Wimmer, 2008; see De Jong and Van der Leij, 2003 for similar results in a Dutch sample). A similar developmental pattern has been found for reading comprehension skills (Klicpera et al., 2006). Hence, targeted interventions are needed to remediate reading difficulties and to counter the negative developmental trajectories.

In this study, we examined a syllable-based word recognition intervention for German fourth graders with reading difficulties. We administered a 16-session training of basic word reading skills to these children and tested the effects of this training on the efficiency of written word recognition and the potential transfer effects on reading comprehension skills at the text level. We define the efficiency of word recognition as fast and accurate recognition of written words. As such, the efficiency of word recognition is an essential aspect of fluent reading and a prerequisite of good reading comprehension (Perfetti, 1985). In what follows, we will argue that a syllable-based intervention has a strong potential to foster poor readers' reading skills. We begin with a discussion of reading interventions that work for struggling readers at the end of primary school. Then, we elaborate on the relevance of word reading fluency for reading comprehension and argue the benefits of focusing on syllables in word recognition training.

An ongoing debate persists as to which interventions are the most effective in fostering basic reading skills at the word level for struggling readers at the end of primary school. The most effective technique to train word recognition skills is phonics instruction, which is based on the alphabetic principle and aims at strengthening systematic associations between graphemes and phonemes. Several meta-analyses report small to medium effects of phonics programs on the accuracy of word decoding ( $d = 0.36$  in Grade 6–12, Edmonds et al., 2009;  $d = 0.49$  in Grade 2–6, Ehri et al., 2001;  $d = 0.26$  in kindergarten to Grade 7, Suggate, 2014) or on aggregated outcome variables of accuracy, speed, and comprehension ( $g = 0.32$  in children and adolescents, Galuschka et al., 2014). However, the effectiveness of phonics instruction seems to depend on grade level. According to a meta-analysis by Suggate (2010), the overall effect size (word reading and text comprehension) of phonics interventions (compared to other types of reading instruction) decreases during primary school education. Phonics instruction seems to work best for struggling readers in Grade 1 (Ehri et al., 2001). Starting at Grade 2, interventions focusing on comprehension become more effective. The basic idea of comprehension instruction is to develop self-regulated meaning making from texts using specific techniques such as question generation or summarizing (Edmonds et al., 2009). These interventions explicitly target on higher-order skills of integration and comprehension and waive teaching phonem-code strategies. Nevertheless, some studies have also found positive side-effects on accuracy and fluency of word reading (Suggate, 2014; Potocki et al., 2015; for effects on fluency, see Müller et al., 2015a).

Despite their potential in fostering reading fluency at the word level, comprehension interventions also require efficient word recognition skills to be effective (Müller et al., 2015b). Comprehension usually involves the teaching of resource-demanding reading strategies, which can be applied successfully only when children need to spend little cognitive effort for recognizing words (cf. lexical quality hypotheses, Perfetti and Hart, 2002). Children who learn to read a consistent language (such as German, Dutch, or Finnish) usually develop from a non-reader to an accurate decoder within the first school year. However, a broad and stable range in decoding fluency exists

even in fourth graders whose word-reading accuracy is close to ceiling (cf. Landerl and Wimmer, 2008). A strong connection between fluency of word recognition and reading comprehension has been observed in students from first to fourth grade ( $r = 0.81$  in a sample of Finnish first and second graders, Torppa et al., 2007;  $r = 0.67$  in a sample of German third and fourth graders, Richter et al., 2017). Knoepke et al. (2014) investigated the relationship between visual word recognition processes (i.e., phonological recoding and orthographical decoding) and reading comprehension in German children in Grade 2 to Grade 4. In this study, the efficiency of both types of word recognition processes was a significant and unique predictor of reading comprehension, whereas their relative weight did not change across grade levels. That is, even in Grade 4 a strong relationship persisted between word recognition skills and reading comprehension. This relationship was strongest for orthographic decoding.

Achieving efficient orthographical decoding skills represents an important footstep in the development of reading fluency. Beginning readers rely primarily on phonological recoding, because most written words are unknown for them. More experienced readers, in contrast, read most words holistically via orthographical decoding by mapping (sub)lexical units or whole word forms directly on to their lexical entries (Frith, 1986; Ehri, 2005). However, some students miss this step in routinization and remain at the alphabetic stage of phonological letter-by-letter recoding. Research on dyslexic primary students in transparent orthographies from Grade 2 onward hint on large word length effects (i.e., impeding effects of the number of letters, Hautala et al., 2012) during word and non-word reading. These results indicate that these children still primarily rely on phonological recoding (Zoccolotti et al., 2005; Martens and de Jong, 2006; cf. Hautala et al., 2012). As a consequence, reading is slow and disfluent and associated with reading comprehension difficulties (Torppa et al., 2007). In sum, given the crucial importance of efficient orthographical decoding skills in reading development, constructing reading interventions targeted at developing these skills is strongly advisable.

According to theoretical models of reading development, children move from slow letter-by-letter decoding to the extraction of units that are larger than phonemes (cf. consolidated alphabetic phase, Ehri, 2005). Empirical studies suggest the importance of the syllable as the sublexical unit that bridges phonology and lexical entries (Hautala et al., 2012) through which word recognition fluency is facilitated. Colé et al. (1999) presented a target syllable followed by a word to French-speaking first graders. The participants' task was to respond when the syllable appeared at the beginning of the word. At the end of Grade 1, only children with high scores in word reading fluency showed significant faster response times when the syllable matched the word. Disfluent children, in contrast, did not show a syllable compatibility effect. Häikiö et al. (2015, 2016) repeatedly found that marking syllables via hyphenation was beneficial for Finnish second graders with poor comprehension skills. In contrast, word recognition slows down significantly in children with good comprehension skills in the hyphenation condition, indicating that good readers already mastered reading with orthographic comparisons and probably accessed more than



one syllable simultaneously (Grainger and Ziegler, 2011). Poor readers, in contrast, seem to experience difficulties in recoding larger chunks of letters. This observation is in accordance with results of Scheerer-Neumann (1981) who presented lists of pseudowords consisting of syllables that also appear in real German words. The pseudowords were presented either with or without graphical syllable segmentation. Poor readers achieved better results in the segmentation condition, whereas good readers showed the same accuracy in both conditions. These studies provide evidence for individual differences in syllable segmentation in beginning readers. Poor readers seem to experience difficulties in recoding syllabic units and reading holistically. Consequently, these readers' word recognition is inefficient and uses a large amount of cognitive resources, which slows down reading.

In reading intervention studies, repeated reading of multi-letter consonant clusters (Hintikka et al., 2008; Huemer et al., 2008), of frequent syllables (Wentink et al., 1997; Bhattacharya and Ehri, 2004; Heikkilä et al., 2013), and of infrequent syllables (Huemer et al., 2010) has been shown to increase the accuracy and fluency of word recognition in children who had received at least 2 years of regular reading instruction. For transparent orthographies with clear syllabic structure, the effect in promoting reading speed is most pronounced for infrequent syllables (Finnish: Huemer et al., 2010; Heikkilä et al., 2013). In German, however, the syllabic structure is shallow and ambiguous (Seymour et al., 2003; Ziegler and Goswami, 2005) making it less easy for developing readers to extract syllabic units reliably. Thus, we assumed an intervention based on the most frequent syllables to be effective in promoting poor readers' reading fluency. Given the consistency of German orthography, even poor readers are likely to acquire the competence to recognize German words accurately via phonological recoding, but their reading will be slow. An intervention based on extensive practice in syllable reading might help poor readers to learn the redundancy and the regularities of letters and words, which can be used for orthographic decoding (Ehri, 2005). We hypothesized that the mental representations of syllables and frequent words consisting of these syllables would be strengthened by the practice in reading materials based on the most frequent syllables. One syllable is contained in several words and can be presented in different positions of the word. Thus, the extensive practice of syllable recognition and segmentation to ameliorate word reading fluency among poor readers might be a promising intervention.

The aim of the current study was to examine effects of a newly developed syllable-based reading intervention for poor German-speaking readers in Grade 4. The theory behind syllabic reading predicts that the intervention should strengthen orthographic decoding processes as an indicator for word reading fluency and, indirectly, even promote reading comprehension (Perfetti and Hart, 2002). Another exploratory aim of the study was to investigate differences between poor and good reading fourth graders immediately after the intervention to assess the degree of improvement of poor readers compared to children without reading difficulties who receive no extracurricular reading training.

## MATERIALS AND METHODS

### Design and Procedure

The study followed an experimental pre/post-test design with randomization at the class level. The data were collected as part of a longitudinal study that examined the effects of different reading interventions in primary school.

We first screened students with two standardized computer-based German-speaking reading tests. Subtests of the ProDi-L (Richter et al., 2012, 2017) were used to capture reading skills at the word level (phonological recoding, orthographic decoding, and access to word meanings) and a subtest of the ELFE (Lenhard and Schneider, 2006) to assess reading comprehension at the text level. We selected children with poor reading skills as participants. Poor reading skills were operationally defined as scores below percentile rank 50 on the class norms of both word recognition (mean composite score of the three ProDi-L subtests) and reading comprehension. Children were clustered in groups of four to six and the groups were randomly allocated at the class level to either the treatment or the control condition.

The nine groups in the treatment condition received the intervention between pre- and post-test. Student assistants (prospective teachers or psychology undergraduates) conducted the 16 treatment sessions, each session lasting 45 min. The training sessions occurred in addition to regular school curriculum twice a week. The efficiency of children's reading processes was assessed again after the final training session with the ProDi-L, and reading comprehension was again assessed with the ELFE. The control condition was a wait-list group. Thus, the seven groups assigned to the control condition received a reading intervention after the post-test.

### Participants

The participating poor readers were 75 fourth graders from nine primary schools (23 classes). Of these, 43 children were allocated to the treatment condition and the remaining 32 to the control condition. The study was conducted in Giessen and Kassel (Germany). For organizational reasons, the treatment took place in Kassel and the control condition was divided between Giessen and Kassel. The average age of the participants was 10;13 years ( $SD = 1$  year) and the proportion of boys and girls was nearly equal in both treatment condition,  $\chi^2(1, N = 75) = 2.34$ ,  $ns$  (see Table 1). The mean Z-values of the word recognition and

**TABLE 1 |** Characteristics of the sample and mean Z-values of word recognition skills and reading comprehension at pretest by treatment condition (compared to class norms).

	Treatment group	Control group
<i>N</i>	43	32
Number of females	21	10
Word recognition (ProDi-L)		
Phonological recoding <i>M</i> ( <i>SD</i> )	−1.11 (0.86)	−0.81 (0.99)
Orthographical decoding <i>M</i> ( <i>SD</i> )	−1.31 (0.62)	−1.22 (0.72)
Access to word meaning <i>M</i> ( <i>SD</i> )	−1.23 (0.60)	−1.12 (0.66)
Reading comprehension (ELFE) <i>M</i> ( <i>SD</i> )	−0.96 (0.59)	−1.07 (0.64)

reading comprehension skills at pretest were below the average in both treatment conditions compared to the class norms (see **Table 1**).

A group of 44 good readers (percentile rank above 50 on the class norms of both word recognition and reading comprehension, all from Kassel) was also tested at pre- and post-test to explore differences between poor and good readers at post-test on both outcome measures.

## Measured Variables

### Assessment of Fluency of Word Recognition

We used a lexical decision task, the subtest orthographical decoding of the instrument ProDi-L (Richter et al., 2012, 2017), to assess the fluency of single-word reading. The children were required to decide whether a string of letters was a real word or a pseudoword. The 16 items, half of which were real German words and the other half pseudowords (orthographically and phonologically legal), varied systematically in length, frequency, and the number of orthographical neighbors. We used parallel versions of the subtest at pre- and post-test. The test recorded accuracy and latency of yes/no responses (provided with two response keys). An integrated test score was calculated as an indicator of the word recognition fluency. Therefore, a quotient of accuracy (mean number of correct responses) and response time (mean response time of the logarithmically transformed response times of all items when at least three items per scale had valid responses) was calculated. Thus, a high score indicates that a reader was faster and more accurate in word recognition fluency. The test-retest reliability was  $r = 0.25$  (computed as the correlation of the pre- and post-test measures in the control group).

### Measurement of Reading Comprehension

Reading comprehension skills were assessed with the subtest text comprehension of the ELFE 1–6 (Lenhard and Schneider, 2006). Children were presented with 20 short texts and were asked to answer questions concerning the content of each text by choosing one of four multiple-choice items. The test score is the sum of correct responses. The same 20 texts were presented at pre- and post-test in randomized order. The test-retest reliability was  $r = 0.55$  (within the control group).

## Intervention: Syllable-Based Word Recognition Training

We designed the materials and the standardized manual of the intervention in cooperation with a learning therapist. All word materials were systematically selected based on the 500 most frequent German syllables in texts typically read by 9–12 year-old children (cf. data base childLex, Schroeder et al., 2015). The exercises included analyzing the syllabic structure of words by marking syllables with arcs during reading, finding the vowel nucleus within each syllable, combining prefixes and stems, and reading words aloud syllable-by-syllable. Special consideration was given to accurate phonological pronunciation of consonant clusters. Word recognition was first trained for single words with a regular spelling and a maximum length of four syllables. In advanced phases of the training, the complexity of materials

increased to irregular words with up to eight syllables, and the materials included sentences and short texts. Several games were used to motivate children to read syllables and words as fast as they could, for instance, a kind of flash card reading (i.e., words were presented very briefly on cards in order to necessitate fast decoding, Wentink et al., 1997), games such as “syllable race” (i.e., moving a game character on a board according to the number of syllables in a word) and “syllable jump” (i.e., jumping on syllable cards on the floor in the order of the syllables in a word presented orally), or detecting two-syllabic words while showing cards with syllables on two stacks. The rationale behind the exercises was to strengthen the mental representations of syllables and orthographic representations that consist of these syllables, resulting in accuracy and speed improvement of word recognition.

The intervention was initially designed for 24 sessions. However, we found that after 3 weeks of intervention the exercises could be implemented faster than expected. Consequently, the number of sessions was reduced to 16 by combining two successive sessions within one.

## RESULTS

### Data Analysis

Given the hierarchical structure of our data (students nested within classes nested within schools), we first estimated the intra-class correlations (ICCs) with random intercept multilevel models (Raudenbush and Bryk, 2002). The ICCs for the fluency of word recognition ( $\rho = 0.10$ ) and reading comprehension ( $\rho = 0.52$ ) indicated clustering effects in the data. Thus, we ran multilevel regression models with intercepts randomly varying between schools and classes. All models were estimated with the software package *lme4* for R (Bates et al., 2015). Significance level was set at 0.05, one-tailed (we tested directed hypotheses). Descriptive statistics and intercorrelations for all variables can be found in **Table 2**.

Separate multilevel regression models were estimated for fluency of word recognition and reading comprehension measures with listwise deletion of missing data. The dummy-coded treatment condition (with the control group as reference category) was used as predictor, fluency and reading comprehension at post-test were used as outcome variables. The pretest score corresponding to the outcome variable was entered as a *z*-standardized covariate to control for pre-training differences. Given the significant pre-training difference in fluency of word recognition between participants in Kassel ( $M = 386.81$ ,  $SD = 47.54$ ) and Giessen ( $M = 427.23$ ,  $SD = 43.12$ ),  $F(1,73) = 8.98$ ,  $p < 0.01$ ,  $\eta^2 = 0.11$ , the city where the data were collected was also entered as a dummy-coded predictor. All predictors were entered simultaneously into the models.

The visual inspection of standardized residuals versus unstandardized predicted values revealed that the assumptions of linearity, normality, and homoscedasticity were not violated in any of the models. The assumption of the independence of residuals was also supported (Cohen et al., 2003, Chap. 4 and 10).

In the two models, two to three cases that deviated more than 2.5 standard deviations from the mean of the residuals were excluded from the analysis (Baayen, 2008, Chap. 7).

## Effects on Fluency of Word Recognition

The results for average effects of the syllable-based intervention on the poor readers' reading skills compared to the control group are shown in **Table 3**. The multilevel regression revealed a significant treatment effect for word recognition fluency [ $\beta = 51.51$ ,  $t(60) = 3.4$ ,  $p < 0.05$ ], which is illustrated in **Figure 1**. The effect size for the average treatment effect (computed as the difference between the adjusted means divided by the standard deviation of the outcome variable of the control group, cf. Mayer et al., 2014) indicated a strong effect ( $ES = 0.82$ ). The adjusted means for the fluency of word recognition at post-test (estimated with the R-package *lmerTest*, Kuznetsova et al., 2015) were considerably higher in the treatment condition ( $M_{\text{adjust}} = 467.76$ ,  $SD = 43.35$ ) than in the control condition ( $M_{\text{adjust}} = 416.24$ ,  $SD = 63.01$ ).

## Effects on Text-Based Reading Comprehension

No significant effect was found between the treatment and the control group at post-test with reading comprehension as the outcome variable (see **Figure 2**). Although the reading

comprehension scores in the treatment condition were higher than in the control condition, the comparison did not provide support for the transfer effects of the syllable-based reading intervention on reading comprehension at the text level. In this context, it is important to note that reading comprehension skills of the poor readers showed high stability from pre- to post-test, which might have prevented establishing a treatment effect.

## Comparison of Good and Poor Readers at Post-test

The estimates for the comparisons of the two groups of poor readers with the untreated good readers are shown in **Table 4**. The comparison between the treatment condition and the good readers revealed a significant negative effect on word recognition fluency [ $\beta = -49.61$ ,  $t(74) = -2.88$ ,  $p < 0.01$ ], indicating that poor readers still differed significantly from good readers in fluency ( $M = 501.82$ ,  $SD = 32.55$ ) after the intervention. However, an even larger difference emerged for the poor readers in the control condition [ $\beta = -80.20$ ,  $t(64) = -3.21$ ,  $p < 0.01$ ].

The results for reading comprehension are somewhat surprising. There was no significant effect on the comparison of the treatment group with good readers [ $\beta = 1.53$ ,  $t(70) = 1.28$ ,  $ns$ ]. Poor readers in the treatment group reached the same level of reading comprehension as good readers. In contrast, the comprehension of the poor readers in the control condition

**TABLE 2 |** Summary of intercorrelations, means, and standard deviations for all variables by treatment condition (treatment group below the diagonal, control group above the diagonal).

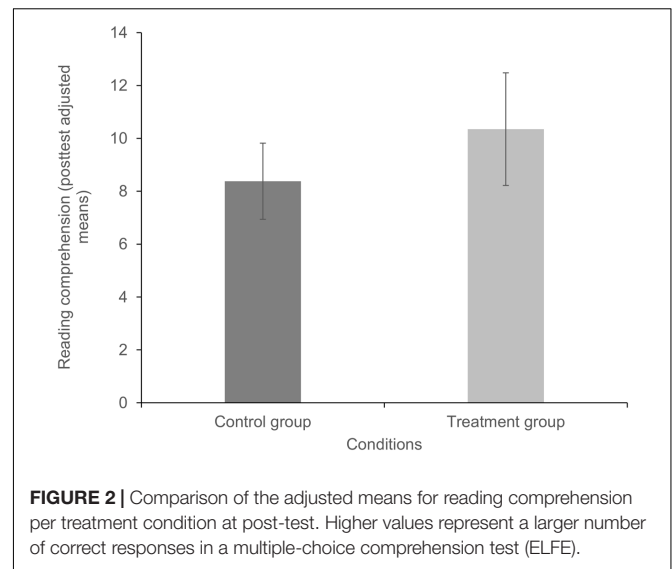
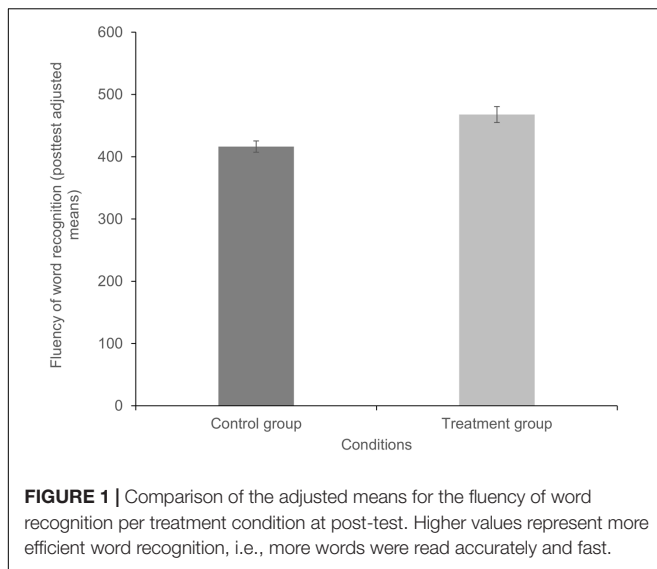
	1	2	3	4	<i>n</i>	<i>M</i>	<i>SD</i>
(1) Fluency of word recognition t1	—	0.25	0.49*	0.23	32	398.29	51.32
(2) Fluency of word recognition t2	−0.13	—	0.49*	0.32	30	414.35	76.86
(3) Reading comprehension t1	0.14	0.51*	—	0.55*	32	6.94	2.83
(4) Reading comprehension t2	0.18	0.37*	0.80**	—	28	8.04	4.56
<i>N</i>	43	37	43	37			
<i>M</i>	392.37	439.24	7.33	10.68			
<i>SD</i>	47.98	44.78	2.67	4.23			

Fluency of word recognition = quotient of accuracy and reaction time, Range = 0–547.67 (Richter et al., 2012); Reading comprehension = sum of correct responses, Range = 0–20 (Lenhard and Schneider, 2006); t1 = pretest; t2 = post-test. Means and standard deviations for the treatment group ( $N = 43$ ) are presented in the rows at the bottom of the table, for the control group ( $N = 32$ ) in the columns on the right side of the table. \* $p < 0.05$ ; \*\* $p < 0.01$ .

**TABLE 3 |** Fixed effects and variance components for the multilevel analyses with fluency of word recognition and reading comprehension as outcome variables, treatment condition as predictor, and corresponding pretest scores and city as covariates.

Parameter	Fluency of word recognition	Reading comprehension
	<i>B</i> ( <i>SE</i> )	<i>B</i> ( <i>SE</i> )
<b>Fixed effects</b>		
Intercept	446.79 (14.18)**	8.59 (1.85)*
Treatment (1) vs. Control Condition (0)	51.51 (15.17)*	1.96 (2.71)
Pretest score (z-standardized)	6.56 (6.72)	2.59 (0.31)**
City (0 = Giessen, 1 = Kassel)	−61.09 (20.21)*	−0.41 (2.90)
<b>Variance components</b>		
Classes within schools	0.00 (0.00)	0.00 (0.00)
Schools	0.00 (0.00)	9.04 (3.01)

\* $p < 0.05$ ; \*\* $p < 0.01$  (one-tailed).



differed significantly from the good readers [ $\beta = -6.47$ ,  $t(61) = -3.67$ ,  $p < 0.01$ ].

## DISCUSSION

The aim of this study was to investigate the effects of a syllable-based reading intervention for poor readers in Grade 4. We assumed that the intervention would lead to gains in word reading fluency and reading comprehension.

The results of this study support primarily the first prediction. The syllable-based training had a significant and strong effect on the fluency of recognizing single words. The effect was similar in size compared to earlier studies on syllable reading interventions in orthographies with unambiguous syllable boundaries (Huemer et al., 2010). This result is encouraging given the complex syllabic structure in German. To the best of our knowledge, only

one other intervention study with German-speaking primary school children investigated a training based on syllables as the unit of analysis (Ritter, 2010). In this study, a small group ( $N = 48$ ) of third and fourth graders with poor reading skills was divided into a treatment group, placebo control group, and wait-list control group. The treatment group received an 18-session training of syllable and morpheme segmentation in one-to-one sessions. The results hint on improvements in word reading fluency and accuracy. The treatment effect of our study strengthens and expands these results. Our material was designed systematically based on the 500 most frequent written syllables for 9 to 12 year-olds, and the focus was on reading those syllables (i.e., frequent words consisting of these syllables on different positions). The intervention was implemented in small groups with a heterogeneous sample of relatively poor readers. We found that poor readers in the treatment condition outperformed same-skilled children in the control group at the post-test in reading

**TABLE 4 |** Fixed effects and variance components for the multilevel analyses with fluency of word recognition and reading comprehension as outcome variables, treatment condition vs. untreated good readers as predictor, and corresponding pretest scores and city as covariates.

Parameter	Treatment group		Control group	
	Fluency of word recognition	Reading comprehension	Fluency of word recognition	Reading comprehension
	<i>B (SE)</i>	<i>B (SE)</i>	<i>B (SE)</i>	<i>B (SE)</i>
<b>Fixed effects</b>				
Intercept	494.77 (9.62)**	13.37 (0.73)**	547.22 (26.59)**	18.13 (2.29)**
Poor Readers (1) vs. Good Readers (0)	−49.61 (17.21)**	1.53 (1.20)	−80.20 (24.96)**	−6.47 (1.76)**
Pretest score (z-standardized)	9.41 (8.59)	4.77 (0.62)**	24.07 (10.28)*	2.02 (0.68)**
City (0 = Giessen, 1 = Kassel)			−60.07 (22.61)*	−2.55 (2.38)
<b>Variance components</b>				
Classes within schools	74.95 (8.66)	0.25 (0.50)	490.60 (22.15)	0.00 (0.00)
Schools	0.00 (0.00)	0.56 (0.75)	0.00 (0.00)	8.77 (2.96)

The city was only included in the comparison of readers in the control condition and the good readers because the poor readers in the treatment condition and the good readers were all participants from Kassel. \* $p < 0.05$ ; \*\* $p < 0.01$  (two-tailed).



fluency and that their deficits compared to the good readers' reading fluency was reduced. These results are promising. One explanation for the strong effect size might be the composition of the sample of poor readers who received the treatment in our study. Meta-analytic results suggest that children with less severe reading impairments (operationalized as below-average readers) experience greater gains from reading interventions compared to children with severe deficits (operationalized as more than 1 standard deviation below the average; Suggate, 2010; Heikkilä et al., 2013; Galuschka et al., 2014).

In contrast to the fluency of word recognition, text-level reading comprehension was not significantly improved by the training. The result that a syllable-based training only led to specific improvements in word reading skills but lack effects on comprehension scores with older poor readers was already found in other samples, for example with struggling 13 year-old French readers (Potocki et al., 2015). Interestingly, most of the previous studies evaluating syllabic reading interventions did not examine effects on reading comprehension (Wentink et al., 1997; Huemer et al., 2010; Heikkilä et al., 2013). An intervention study by McCandliss et al. (2003) suggests that reading comprehension of poor readers above Grade 1 benefits from word recognition training if word reading is embedded in texts and if the participating children are repeatedly asked to elaborate on the meaning of those texts. Although words were read in the context of short texts in the second half of the intervention presented here, comprehension was not in the focus of the training. For example, exercises involving word meanings or comprehension of words in their linguistic contexts were not included in the training, explaining the lack of effects on comprehension. Furthermore, a longitudinal study with struggling Finnish-speaking readers in Grade 1 and 2, Torppa et al. (2007) found that reading comprehension developed rapidly only after the word recognition skills had reached a sufficient level (cf. Juul et al., 2014). Thus, it seems worthwhile to investigate which level of fluent syllable reading is required to achieve transfer effects on higher-order comprehension.

Another interpretation of the results for reading comprehension is fed by the observation that comprehension scores of poor readers who received the training did not differ significantly from those of the good readers in the sample. In contrast, the post-test reading comprehension scores of the poor readers in the control group still differed significantly from the good readers. This pattern of results suggests that the syllable-based training might indeed have had a positive effect on comprehension. The lack of a significant effect compared to the control group might have been due to a lack of power that was not sufficient to detect small to medium treatment effects.

Despite these open questions, the result of the specific improvement in word reading fluency among German-speaking

fourth graders with poor word recognition and reading comprehension skills is encouraging against the background of the persistence of individual differences in reading fluency (De Jong and Van der Leij, 2003; Landerl and Wimmer, 2008) and the relevance of fluent reading for sufficient reading skills in transparent orthographies. A short, but focused intervention making frequent syllables salient as a unit of recognition seems to have the potential to raise the fluency of word recognitions even in poor readers who have already received at least 3 years of regular reading instruction.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of "Ministry of Education and Cultural Affairs, Hesse, Germany (Hessisches Kultusministerium)" with written informed consent from all subjects. The parents of all subjects gave written informed consent in accordance with the Declaration of Helsinki. Principals of the schools participating in the study gave written consent after the school conference (i.e., the majority of teachers agreed to realize the study in their school).

The protocol was approved by the "Ministry of Education and Cultural Affairs, Hesse" (cf. Education Act of Hesse, section 84).

## AUTHOR CONTRIBUTIONS

The study reported here was realized in cooperation between the authors. The material and manual of the intervention were designed by BM and TR in cooperation with a learning therapist. Data collection took place in Kassel (organized by BM and TR) and Giessen (organized by SK and ME). Statistical analyses were done by BM, PK, and TR.

## ACKNOWLEDGMENTS

The research reported in this article was supported by the German Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF, grant 01GJ1004) and the University of Würzburg in the funding program Open Access Publishing. We would like to thank Gabriele Otterbein-Gutsche for her help in conceptualizing the treatment and all the students, their teachers, and many student assistants for their collaboration in the study. Researchers who would like to inspect the items of the ProDi-L test or the materials of the trainings used in this study are invited to send an e-mail to the first or second author.

## REFERENCES

- Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge, Cambridge University Press.
- Bates, D., Maechler, M., and Bolker, B. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

- Bhattacharya, A., and Ehri, L. (2004). Graphosyllabic analysis helps adolescent struggling readers read and spell words. *J. Learn. Disabil.* 37, 331–348. doi: 10.1177/00222194040370040501
- Bos, W., Bremerich-Vos, A., Tarelli, I., and Valtin, R. (2012). "Lesekompetenzen im internationalen Vergleich [International comparison of reading skills]," in *IGLU 2011: Lesekompetenzen von Grundschulkindern im internationalen Vergleich*

- [IGLU 2011: International Comparison of Reading Skills in Primary School], eds W. Bos I, A. Bremerich-Vos, and K. Schwippert (Münster: Waxmann), 91–136.
- Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences*. Mahwah, NJ: Erlbaum.
- Colé, P., Magnan, A., and Grainger, J. (1999). Syllable-sized units in visual word recognition: evidence from skilled and beginning readers of French. *Appl. Psycholinguist.* 20, 507–532. doi: 10.1017/S0142716499004038
- De Jong, P. F., and Van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *J. Educ. Psychol.* 95, 22–40. doi: 10.1037/0022-0663.95.1.22
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., et al. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Rev. Educ. Res.* 79, 262–300. doi: 10.3102/0034654308325998
- Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Sci. Stud. Read.* 9, 167–188. doi: 10.1207/s1532799xssr0902\_4
- Ehri, L. C., Nunes, S. R., Stahl, S. A., and Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: evidence from the national reading panel's meta-analysis. *Rev. Educ. Res.* 71, 393–447. doi: 10.3102/00346543071003393
- Frith, U. (1986). A developmental framework for developmental dyslexia. *Ann. Dyslexia* 36, 69–81. doi: 10.1007/BF02648022
- Galuschka, K., Ise, E., Krick, K., and Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. *PLOS ONE* 9:e89900. doi: 10.1371/journal.pone.0089900
- Grainger, J., and Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Front. Psychol.* 2:54. doi: 10.3389/fpsyg.2011.00054
- Häikiö, T., Bertram, R., and Hyönä, J. (2016). The hyphen as a syllabification cue in reading bisyllabic and multisyllabic words among Finnish 1st and 2nd graders. *Read. Writ.* 29, 159–182. doi: 10.1007/s11145-015-9584-x
- Häikiö, T., Hyönä, J., and Bertram, R. (2015). The role of syllables in word recognition among beginning Finnish readers: evidence from eye movements during reading. *J. Cogn. Psychol.* 27, 562–577. doi: 10.1080/20445911.2014.982126
- Hautala, J., Aro, M., Eklund, K., Lerkkanen, M.-K., and Lyytinen, H. (2012). The role of letters and syllables in typical and dysfluent reading in a transparent orthography. *Read. Writ.* 26, 845–864. doi: 10.1007/s11145-012-9394-3
- Heikkilä, R., Aro, M., Närhi, V., and Westerholm, J. (2013). Does training in syllable recognition improve reading speed? A computer-based trial with poor readers from second and third grade. *Sci. Stud. Read.* 17, 398–414. doi: 10.1080/10888438.2012.753452
- Hintikka, S., Landerl, K., Aro, M., and Lyytinen, H. (2008). Training reading fluency: is it important to practice reading aloud and is generalization possible? *Ann. Dyslexia* 58, 59–79. doi: 10.1007/s11881-008-0012-7
- Huemer, S., Aro, M., Landerl, K., and Lyytinen, H. (2010). Repeated reading of syllables among finnish-speaking children with poor reading skills. *Sci. Stud. Read.* 14, 317–340. doi: 10.1080/10888430903150659
- Huemer, S., Landerl, K., Aro, M., and Lyytinen, H. (2008). Training reading fluency among poor readers of German: many ways to the goal? *Ann. Dyslexia* 58, 115–137. doi: 10.1007/s11881-008-0017-2
- Juul, H., Poulson, M., and Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *J. Educ. Psychol.* 106, 1096–1106. doi: 10.1037/a0037100
- Klicpera, C., Schabmann, A., and Gasteiger-Klicpera, B. (2006). Die mittelfristige Entwicklung von Schülern mit Teilleistungsschwierigkeiten im Bereich der Lese- und Rechtschreibschwierigkeiten [The development of reading and spelling skills of students with learning disabilities during elementary school]. *Kindheit Entwicklung* 15, 216–227. doi: 10.1026/0942-5403.15.4.216
- Knoepke, J., Richter, T., Isberner, M.-B., Naumann, J., and Neeb, Y. (2014). Phonological recoding, orthographic decoding, and comprehension skills during reading acquisition. *Z. Erziehungswiss.* 17, 447–472. doi: 10.1007/s11618-014-0560-z
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2015). *lmerTest: Tests in Linear Mixed Effects Models. R Package Version 2.0-33*. Available at: <http://CRAN.R-project.org/package=lmerTest>
- Landerl, K., and Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: an 8-year follow-up. *J. Educ. Psychol.* 100, 150–161. doi: 10.1037/0022-0663.100.1.150
- Lenhard, W., and Schneider, W. (2006). *ELFE 1-6: Ein Leseverständnistest für Erst- bis Sechstklässler [ELFE 1-6: A Reading Comprehension Test for Children in Grades One to Six]*. Göttingen: Hogrefe.
- Martens, V. E. G., and de Jong, P. F. (2006). The effect of visual word features on the acquisition of orthographic knowledge. *J. Exp. Child Psychol.* 93, 337–356. doi: 10.1016/j.jecp.2005.11.003
- Mayer, A., Nagengast, B., Fletcher, J., and Steyer, R. (2014). Analyzing average and conditional effects with multigroup multilevel structural equation models. *Front. Psychol.* 5:304. doi: 10.3389/fpsyg.2014.00304
- McCandliss, B., Beck, I. L., Sandak, R., and Perfetti, C. (2003). Focusing attention on decoding for children with poor reading skills: design and preliminary tests of the word building intervention. *Sci. Stud. Read.* 7, 75–104. doi: 10.1207/S1532799XSSR0701\_05
- Müller, B., Mayer, A., Richter, T., Križan, A., Hecht, T., and Ennemoser, M. (2015a). Differential effects of reading trainings on reading processes: a comparison in Grade 2. *Z. Erziehungswiss.* 18, 489–512. doi: 10.1007/s11618-015-0648-0
- Müller, B., Richter, T., Križan, A., Hecht, T., and Ennemoser, M. (2015b). Word recognition skills moderate the effectiveness of reading strategy training in Grade 2. *Learn. Individ. Differ.* 40, 55–62. doi: 10.1016/j.lindif.2015.05.001
- Perfetti, C. A. (1985). *Reading Ability*. New York, NY: Oxford University Press.
- 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), 189–213.
- Potocki, A., Magnan, A., and Ecalle, J. (2015). Computerized trainings in four groups of struggling readers: specific effects on word reading and comprehension. *Res. Dev. Disabil.* 4, 83–92. doi: 10.1016/j.ridd.2015.07.016
- Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd Edn. Thousand Oaks, CA: Sage.
- Richter, T., Isberner, M.-B., Naumann, J., and Kutzner, Y. (2012). Prozessbezogene Diagnostik von Lesefähigkeiten bei Grundschulkindern [Process-oriented assessment of reading skills in primary school children]. *Z. Pädagog. Psychol.* 26, 313–331. doi: 10.1024/1010-0652/a000079
- Richter, T., Naumann, J., Isberner, M.-B., Neeb, Y., and Knoepke, J. (2017). *ProDi-L: Prozessbasierte Diagnostik von Lesefähigkeiten bei Grundschulkindern [ProDi-L: Process-Oriented Assessment of Reading Skills in Primary School Children; Computerized Test]*. Göttingen: Hogrefe.
- Ritter, C. (2010). *Empirische Evaluation eines Lesetrainings auf Silbenbasis zur Förderung der basalen Lesefähigkeit von Grundschulkindern mit Leseschwierigkeiten. Abschlussbericht an die DFG [Evaluation of a Syllable-Based Reading Intervention For Primary School Children. Final Report]*. Potsdam: Universität Potsdam.
- Scheerer-Neumann, G. (1981). The utilization of intraword structure in poor readers: Experimental evidence and a training program. *Psychol. Res.* 43, 155–178. doi: 10.1007/BF00309827
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., and Kliegl, R. (2015). childLex: a lexical database of German read by children. *Behav. Res. Methods* 47, 1085–1094. doi: 10.3758/s13428-014-0528-1
- 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
- Suggate, S. P. (2010). Why what we teach depends on when: grade and reading intervention modality moderate effect size. *Dev. Psychol.* 46, 1556–1579. doi: 10.1037/a0020612
- Suggate, S. P. (2014). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *J. Learn. Disabil.* 49, 1–20. doi: 10.1177/0022219414528540
- Torppa, M., Tolvanen, A., Poikkeus, A.-M., Eklund, K., Lerkkanen, M.-K., Leskinen, E., et al. (2007). Reading development subtypes and their early characteristics. *Ann. Dyslexia* 57, 3–32. doi: 10.1007/s11881-007-0003-0
- Wentink, H. W. M. J., Van Bon, W. H. J., and Schreuder, R. (1997). Training of poor readers' phonological decoding: evidence for syllable-bound processing. *Read. Writ.* 9, 163–192. doi: 10.1023/A:1007921805360
- 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

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., and Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain Lang.* 93, 369–373. doi: 10.1016/j.bandl.2004.10.010

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2017 Müller, Richter, Karageorgos, Krawietz and Ennemoser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# Sensitivity to Stroke Emerges in Kindergartners Reading Chinese Script

Su Li<sup>1</sup> and Li Yin<sup>2\*</sup>

<sup>1</sup> CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China, <sup>2</sup> Center for the Study of Language and Psychology, Department of Foreign Languages and Literatures, Tsinghua University, Beijing, China

## OPEN ACCESS

### Edited by:

Giseli Donadon Germano,  
Universidade Estadual Paulista, Brazil

### Reviewed by:

Mercedes Inda-Caro,  
Universidad de Oviedo, Spain  
Hsu-Wen Huang,  
City University of Hong Kong,  
Hong Kong  
Wei-Lun Chung,  
National Taiwan Normal University,  
Taiwan

### \*Correspondence:

Li Yin  
yinl@tsinghua.edu.cn

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 30 November 2016

**Accepted:** 15 May 2017

**Published:** 02 June 2017

### Citation:

Li S and Yin L (2017) Sensitivity  
to Stroke Emerges in Kindergartners  
Reading Chinese Script.  
Front. Psychol. 8:889.  
doi: 10.3389/fpsyg.2017.00889

To what extent are young children sensitive to individual stroke, the smallest unit of writing in Chinese that carries no phonological or semantic information? The present study examined Chinese kindergartners' sensitivity to stroke and the contribution of reading ability and age to stroke sensitivity. Fifty five children from Beijing, including 28 4-year-olds ( $M_{\text{age}} = 4.55$  years,  $SD = 0.28$ , 16 males) and 29 5-year-olds ( $M_{\text{age}} = 5.58$  years,  $SD = 0.30$ , 14 males), were administered an orthographic matching task and assessed on non-verbal IQ and Chinese word reading. In the orthographic matching task, children were asked to decide whether two items were exactly the same or different in three conditions, with stimuli being correctly written characters (e.g., “分”), stroke-missing or redundant characters (e.g., “𠂇”), and Tibetan alphabets (e.g., “ཨ”), respectively. The stimuli were presented with E-prime 2.0 software and were displayed on a Surface Pro. Children responded by touching the screen and reaction time was used as a measure of processing efficiency. The 5-year-olds but not the 4-year-olds processed correctly written characters more efficiently than stroke-missing/redundant characters, suggesting emergence of stroke sensitivity from age 5. The 4- and 5-year-olds both processed correctly written characters more efficiently than Tibetan alphabets, ruling out the possibility that the 5 year olds' sensitivity to stroke was due to the unusual look of the stimuli. Hierarchical regression analyses showed that Chinese word reading explained 10% additional variance in stroke sensitivity after having statistically controlled for age. Age did not account for additional variance in stroke sensitivity after having considered Chinese word reading. Taken together, findings of this study revealed that despite the visually highly complex nature of Chinese and the fact that individual stroke carries no phonological or semantic information, children develop sensitivity to stroke from age 5 and such sensitivity is significantly associated with reading experience.

**Keywords:** orthographic sensitivity, stroke, early reading, Chinese, kindergartner

## INTRODUCTION

The acquisition of reading skill is a crucial task in child development. Detection of the presence or absence of a stroke in a given character, e.g., “丿” in “禾”, entails a combination of visual skills and orthographic knowledge, both of which play important roles in early reading development (e.g., visual skills: Huang and Hanley, 1995; Ho and Bryant, 1997; Siok and Fletcher, 2001;



Valdois et al., 2004; McBride-Chang et al., 2011; Franceschini et al., 2012; Luo et al., 2013; Van der Leij et al., 2013; orthographic knowledge: Badian, 1994; Cassar and Treiman, 1997; Ho and Bryant, 1997; Lonigan et al., 2000; Cunningham et al., 2001; Li et al., 2006; Wang et al., 2015). The purpose of the current study is to examine to what extent young children are sensitive to individual stroke in a character, the smallest unit of writing in Chinese that carries no phonological or semantic information.

Children develop knowledge about the visual-orthographic characteristics of the writing they are exposed to before receiving formal literacy instruction (e.g., English: Lavine, 1977; Levy et al., 2006; Treiman et al., 2007; Pollo et al., 2009; Chinese: Ho et al., 2003; Luo et al., 2011; Yin and McBride, 2015). For example, 3-year-old English-speaking children accepted Latin letters as writing more often than visually dissimilar symbols such as Chinese characters (Lavine, 1977). Pre-phonological 4-year-olds from Brazil and United States produced spellings that reflect consistent differences in spelling pattern between Portuguese and English, indicating their implicit abstraction of regularities in the prints they are exposed to (Pollo et al., 2009). Chinese 5-year-olds learned to read significantly better when the subcomponents of character stimuli were legally positioned than when they were illegally positioned, when no phonetic cue was available, suggesting their implicit knowledge about the positional regularities in Chinese (Yin and McBride, 2015).

Most of the previous studies, however, examined children's visual-orthographic knowledge at the level of word component, such as letter/letter string in alphabetic languages or character/radical (stroke combination that recurs across characters) in Chinese. Word components of such carry, in varying degrees, information of sound and meaning (although orthography, phonology, and semantics are intrinsically inseparable across writing systems). Few studies have examined whether children are sensitive to the elements within these components, e.g., the stroke “乚” in character “禾”, which carries no phonological or semantic information. The only study that examined young children's awareness of stroke (Luo et al., 2011) asked children to discriminate between Chinese stroke and English letter in isolation, e.g., “乚” and “f”. Stronger evidence for stroke sensitivity, however, should come from testing whether children can detect the removal of a single stroke from (e.g., “乚” in “禾”) or the addition of a single stroke to (e.g., “乚” in “云”) a real character.

In the present study, we investigated whether Chinese 4- and 5-year-old kindergartners develop sensitivity to stroke in Chinese character and how such sensitivity is linked to reading ability and maturation.

Chinese orthography is known for high visual complexity. Words in alphabetic scripts are formed from a limited visual set (e.g., 26 letters in English and 22 in Hebrew), but the 1000s of characters in Chinese represent thousands of visually different stroke configurations. A stroke is a dot or a line written in one continuous movement (Anderson et al., 2013). As in character “热”, a dot can be in various directions (“丶” or “㇀”) and a line can be horizontal (一), vertical (丨), slanting (丿), curved (㇏) or contain a hook at the end (㇚). Around 80% of modern Chinese characters are compound characters composed of radicals, i.e.,

stroke patterns that recur across characters. The semantic radical (e.g., “讠” in “语” and “详”) gives a clue to the character's meaning and the phonetic radical (e.g., “马” in “妈” and “骂”) gives a clue to its pronunciation, though the cuing effect is not consistently reliable. Some radicals appear consistently in a fixed position in the character (e.g., “讠” always on the left, “阝” always on the right, “+” always on the top). About 20% characters are simple characters that are not divisible into sub-components and thus are composed of individual strokes only (e.g., “禾”).

Stroke sensitivity is important in learning to read Chinese. First, sensitivity to the identity and relative position of a stroke helps discriminate between characters that are visually similar but totally different in sound and meaning, which are abundant in Chinese (e.g., 土 [/tu/, soil] and 士 [/shi/, soldier], 未 [/wei/, not] and 末 [/mo/, end], 天 [/tian/, sky] and 夫 [/fu/, husband]). Second, it helps discover the internal structure of compound characters, identify radicals, and distinguish between visually similar radicals (e.g., “𠂇” and “𠂈”, “𠂉” and “𠂊”); such abilities are important for recognition of Chinese characters (Shu et al., 2003).

We hypothesized that despite the high visual complexity of Chinese and the absence of linguistic information in stroke, Chinese kindergartners may develop sensitivity to stroke before formally learning to read, and such sensitivity is significantly linked to reading ability in addition to maturational age. The hypotheses were based on a number of reasons.

First, orthographic processing is at least partly constrained by visual object processing that involves extraction of perceptual features of individual elements to identify visual objects (Grainger et al., 2012). With repeated exposure to letter combinations of manipulated frequency, baboons (non-human primates) learned to discriminate English words from non-words through extracting features of individual letters and their combinations. Characters represent a special class of visual objects exposed to Chinese children. There is evidence that Chinese children begin to pay more attention to the visual form information of words in highly familiar environment from age 5 (Zhao et al., 2014b).

Second, Chinese children seem to develop stronger visual skills than children learning to read alphabetic orthographies due to the visually demanding nature of Chinese orthography (Huang and Hanley, 1994; Demetriou et al., 2005; McBride-Chang et al., 2011). Chinese children showed clear advantage over British children on visual form discrimination skills (Huang and Hanley, 1994), and Chinese kindergartners outperformed Israeli and Spanish peers on task of visual spatial relationships (McBride-Chang et al., 2011). We reason that Chinese children's better visual skills may facilitate stroke processing and may in fact reflect the consequence of stroke processing (Zhou et al., 2014).

Third, Chinese kindergartners have developed quite some knowledge about the visual-orthographic features of Chinese before receiving formal literacy instruction (Ho et al., 2003; Luo et al., 2011; Wang et al., 2015; Yin and McBride, 2015). In a character learning task (Yin and McBride, 2015), the 4-year-olds learned pseudocharacters (character stimuli consisting of real radicals placed in legal positions) and non-characters (character stimuli consisting of real radicals placed in illegal positions) significantly better than random

stroke combinations, suggesting their structural knowledge of characters; the 5-year-olds learned pseudocharacters significantly better than non-characters, reflecting their knowledge of the identities and positional regularities of radicals. Luo et al. (2011) found that Chinese 5-year-olds could discriminate between Chinese stroke and English letter when presented in isolation (e.g., 丿 and f). In the present study, we tested Chinese 4- and 5-year-olds' stroke sensitivity at a much finer level, examining whether children can detect the removal or addition of one stroke from a real character (e.g., “丿” in “禾”, or “丿” in “云”).

Finally, reading ability modulates visual expertise for word processing across orthographies (Burgund et al., 2006; Li et al., 2013; Zhao et al., 2014a; Su et al., 2015). Burgund et al. (2006) found the emergence of letter specific-processing in English is linked to increased reading skill rather than increased age among a group of 6-to-19-year-olds. Zhao et al. (2014a) found that fine neural tuning for visual words, which reflects sensitivity to orthographic regularity, emerged in 7-year-old German-speaking children with high but not low reading ability. Li et al. (2013) found that neural specialization for word processing is significantly influenced by reading experience (indexed by sight vocabulary) in 5- and 6-year-old Chinese kindergartners.

In the present study, we tested a group of 4 and 5-year-old kindergarten children in Beijing who have not received formal literacy instruction. We used an orthographic matching task to tap children's sensitivity to stroke in character. There were three orthographic conditions, using stimuli of correctly written characters, stroke missing or redundant characters, and Tibetan alphabets, respectively. In each condition there were 20 stimulus pairs, half of which were the same and half of which were different. Using a Surface Pro with E-Prime 2.0 software to present stimuli pairs, we asked children to decide whether the two items on the screen were exactly the same or different by touching the corresponding happy or sad face. This design allowed us to analyze children's reaction time as a measure of processing efficiency, which was more objective and reliable than oral reports, as were typically used in previous research with young children.

If children were sensitive to stroke, they would process correctly written characters more efficiently than stroke missing/redundant characters. To rule out the possibility that children's lower efficiency in processing stroke missing/redundant characters is due to foreignness of the stimuli's look (e.g., 禾) rather than sensitivity to the missing/redundant stroke, we added the condition of Tibetan alphabets because Tibetan alphabet looks very different from Chinese character both in terms of the shape of its constituents and the configuration matter of its constituents (e.g., 𑄀). Children who were sensitive to stroke should show lower efficiency in processing both Tibetan alphabets and stroke-missing/redundant characters compared with correctly written characters, whereas children who have not yet developed sensitivity to stroke may show lower efficiency in processing Tibetan alphabets but not stroke-missing/redundant characters compared with correctly written characters.

We also assessed children's Chinese word reading ability. Based on findings from previous studies, we predicted a unique

contribution of reading ability to the emergence of stroke sensitivity. In view of the young age of the participants in the current study (4- and 5-year-olds) and the more demanding nature of the task (processing the smallest unit in almost the most visually complex orthography—Chinese), we might also expect that maturation play an equally important role in the emergence of stroke sensitivity.

## MATERIALS AND METHODS

### Children

Fifty five children from Beijing participated in the study. They were 28 4-year-olds ( $M_{\text{age}} = 4.55$  years,  $SD = 0.28$ , 16 males) and 29 5-year-olds ( $M_{\text{age}} = 5.58$  years,  $SD = 0.30$ , 14 males). All children were native Chinese speakers and had normal vision and no known disorders. The researchers explained the purpose and procedure of the study to children's parents/guardians and obtained consents from all participating children's parents/guardians. Children were allowed to withdraw at any time of the study and their rights and privacy were protected throughout the study according to the American Psychological Association Ethical Principles of Psychologists and Code of Conduct, Including 2010 Amendments<sup>1</sup>.

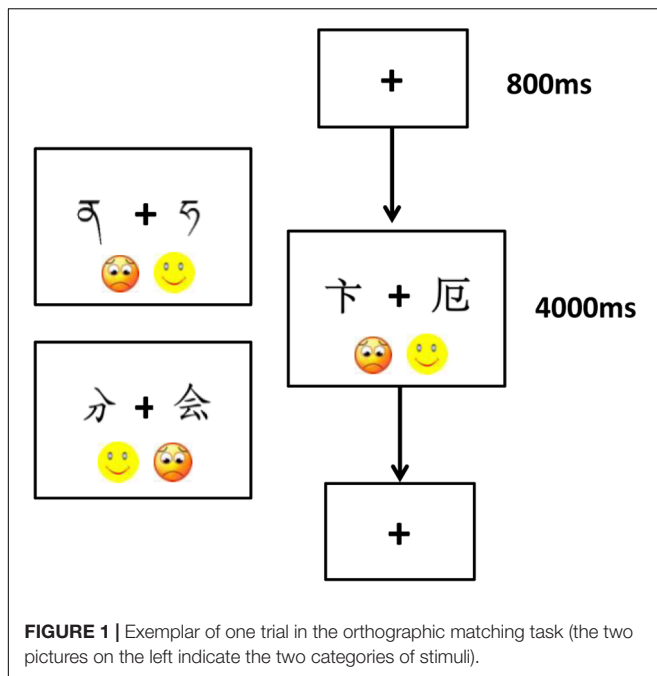
### Materials

The experimental materials included three types of stimuli: correctly written characters, stroke-missing/redundant characters, and Tibetan alphabets. There were 20 items for each type of stimulus. As shown in **Table 1**, the correctly written characters were common characters which included five simple characters, eight left–right structured compound characters, and seven bottom–up structured compound characters. The number of the three structures was based on their distributions in the total number of Chinese characters (Dictionary of Chinese Character Information, 1988, as cited in Shu et al., 2003). The average frequency of the correctly written characters was 55985.5 per 10 millions (range: 22795.5–379917), and the average number of the strokes was 6 (range: 2–10). The stroke-missing/redundant characters were constructed by adding or deleting one stroke from the correctly written characters, with half of them made from adding one stroke to the correctly written characters and the other half made from deleting one stroke from the correctly

<sup>1</sup><http://www.apa.org/ethics/code/index.aspx>

**TABLE 1** | Samples of the three types of stimuli used in the orthographic matching task.

Correctly written characters	Stroke-missing/redundant characters	Tibetan alphabets
禾	禾	𑄀
分	分	𑄁
云	云	𑄂
上	上	𑄃



written characters. The Tibetan alphabet letters were selected randomly from the 36 consonant letters in Tibetan. The average number of the letter strokes was 4 (range: 3–6). While we were primarily concerned with the performance difference between correctly written characters and stroke-missing/redundant characters, we compared performance on correctly written characters and Tibetan alphabet letters as well in order to elucidate the difference between correctly written characters and stroke-missing/redundant characters. All stimuli were presented in 180-point regular script font with the same size of  $300 \times 350$  pixels. All stimuli were presented centrally on the screen, in black against a white background.

## Tasks

### General Cognitive Ability Measurement

We administered the Combined Raven's Test (CRT-Chinese version, 1991) to assess children's IQ. The internal consistency reliability for this task was 0.86.

### Orthographic Matching Task

We used an orthographic matching task to tap children's sensitivity to strokes. There were three orthographic conditions using stimuli of correctly written characters, stroke-missing/redundant characters, and Tibetan alphabets, respectively. As shown in **Figure 1**, in each condition, there were 20 stimulus pairs, half of which were exactly the same and half of which were different. We used E-prime 2.0 software to present the stimuli and used a Surface Pro to display the stimulus pairs. Children responded by touching the screen so that we could obtain their direct response to the stimulus pair. The pairing of stimuli, the left–right positioning of stimulus in the pair, and the order of pair presentation were randomized in each condition. The order of condition was counterbalanced across children.

In each trial (**Figure 1**), following a 800 ms fixation, each stimulus pair was presented until the children made response. Children were instructed to decide whether the two items on the screen were exactly the same or different and to indicate their response as quickly as possible by using their index finger of the right hand to touch the happy face (“😊”) or sad face (“😞”) on the screen. Touching the happy face indicated a “same” response and touching the sad face indicated a “different” response. The left–right positioning of the two faces was counterbalanced across children. Six practice trials were provided before the formal experiments began to ensure that children understood how to perform in the task.

We used a “catching butterfly” play to obtain the basic reaction time of each child. Children were asked to touch as quickly as they could a black butterfly that appeared in the middle of the screen. The play was presented through E-prime 2.0 software on the Surface Pro as well. The presenting time of each butterfly was 2000 ms and the reaction time widows were between 2000 and 8000 ms with 1500 ms as a jitter. The picture of butterfly would disappear after children touched it. The play contained 12 trials including two practice trials.

### Chinese Word Reading Task

Children were asked to read aloud 50 Chinese single-character words presented in order of increasing difficulty (Wang et al., 2015). Testing was discontinued when children failed to read 10 characters consecutively. One point was given for each correctly read item. The maximum score was 50. The internal consistency reliability for this task was 0.96.

## Procedure

Children completed four tasks over a period of 2 weeks in the first semester of the school year. All tasks were administered individually by trained graduate students in a quiet reading room. Half of the children in each age group completed the orthographic matching task, preceded by the basic reaction time measurement, in the first week, and the Raven's test and the Chinese word reading task in the second week; the other half of children completed the Raven's test and the Chinese word reading task in the first week, and the orthographic task preceded by the basic reaction measurement in the second week. The orthographic matching task lasted 20–25 min, with a 5-min break in the middle. The Raven's test and the Chinese word reading task took approximately 10 min each.

## RESULTS

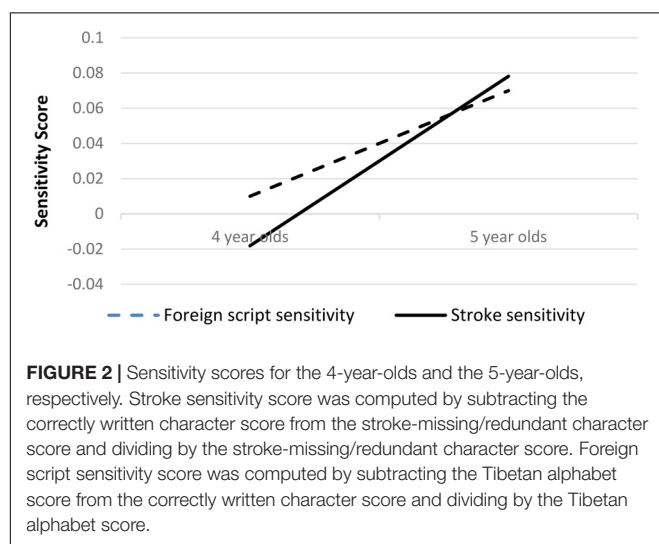
### Descriptive Statistics

**Table 2** shows children's performance (raw score) in each task as a function of age group. In the orthographic matching task, the accuracy rate of the 4-year-olds was 0.96, 1.00, and 0.94 for the correctly written characters, stroke-missing/redundant characters, and Tibetan alphabets, respectively, with no significant difference across conditions,  $F(2,81) = 2.00$ ,  $p = 0.14$ ; similarly, the accuracy rate of the 5-year-olds was 1.00, 0.97, and 0.96 for the correctly written characters,

**TABLE 2 |** Mean (standard deviation) of age, basic reaction time, reaction time in orthographic matching, IQ, and Chinese word reading as a function of age group.

	Age group		Group difference $F(1,55)$
	4 year olds	5 year olds	
<i>N</i>	28	29	
Age (in year)	4.55 (0.27)	5.58 (0.29)	187.52***
Basic reaction time	944.57 (263.80)	741.62 (154.49)	12.58**
Reaction time in orthographic matching			
Correctly written character	2653.10 (519.25)	2130.26 (282.69)	22.50***
Stroke-missing/redundant character	2671.13 (656.37)	2333.87 (396.99)	5.55*
Tibetan alphabet	2755.47 (649.14)	2332.41 (396.68)	8.89**
IQ	10.89 (3.53)	15.00 (3.86)	17.06***
Chinese word reading	8.46 (12.32)	16.17 (14.08)	4.71*

\*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .



stroke-missing/redundant characters, and Tibetan alphabets, respectively, with no significant difference across conditions,  $F(2,84) = 1.22$ ,  $p = 0.30$ . Given that the task was designed to be easy for young children to understand and perform (children were just to decide whether the two stimuli on the screen were same or not at the perceptual level), the high accuracy rate across conditions was expected, indicating that the participants performed well and the data was reliable. We then focused on reaction time of the correct responses as a measure of processing efficiency, i.e., an index of sensitivity, in subsequent analyses. Reactions times that were over two standard deviations away from the mean in a given condition in each group were removed from the data. Overall, 1.8% of the total 4860 trials were removed from the data for this reason.

Following Burgund et al. (2006), to reduce spurious distortions of within-subject differences due to overall group differences (Chapman and Chapman, 1973; Chapman et al., 1994), we computed two sensitivity scores in the following manner for each child. Stroke sensitivity score was computed by subtracting the correctly written character score from the stroke-missing/redundant character score and dividing

by the stroke-missing/redundant character score ([stroke-missing/redundant character score-correctly written character score]/stroke-missing/redundant character score). Foreign script sensitivity score was computed, for purpose of discriminating stroke sensitivity from sensitivity caused by the foreign or unusual look of stimuli, by subtracting the Tibetan alphabet score from the correctly written character score and dividing by the Tibetan alphabet score ([Tibetan alphabet score-correctly written character score]/Tibetan alphabet score). **Figure 2** shows the scores of stroke sensitivity and foreign script sensitivity in each group.

Before analyzing, we checked the normality of distribution of the data. Shapiro–Wilk tests showed that the data of foreign script sensitivity was normally distributed in both age groups,  $p_s > 0.05$ , and the data of stroke sensitivity was normally distributed in the 5-year-olds,  $p = 0.64$ , but not in the 4-year olds,  $p = 0.002$ . Considering the relatively small sample size in the current study and that parametric analysis of transformed data is a better strategy than non-parametric analysis because it appears to be more powerful than the latter (Rasmussen and Dunlap, 1991), we normalized the data of stroke sensitivity and foreign sensitivity using rank-case transformation, which was reported as working better for small sample size in association tests than logarithm and Box-Cox transformations (Goh and Yap, 2009). We used the normal scores obtained using Blom's formula as dependent variables in the following analyses.

## Development of Stroke Sensitivity

Taking basic reaction time as the covariate and age group as the fixed effect, univariate analysis was conducted for stroke sensitivity and foreign script sensitivity, respectively. Homogeneity of the covariate coefficient was checked first. For both stroke sensitivity and foreign script sensitivity, the interaction between age group and basic reaction time was not significant,  $p_s > 0.12$ , indicating that homogeneity of the covariate coefficient across age groups could be assumed in both analyses.

For stroke sensitivity, the 5-year-olds were significantly higher than the 4-year-olds,  $F(1,54) = 6.89$ ,  $p = 0.011$ ,



$\eta_p^2 = 0.11$ . The 5-year-olds processed correctly written characters (e.g., “分”) more efficiently, as reflected by the positive score of stroke sensitivity, than stroke-missing/redundant characters (e.g., “𠂇”), whereas the 4-year-olds did not. For foreign script sensitivity, no significant group difference was found,  $F(1,54) = 0.15$ ,  $p = 0.70$ . The 5 year olds and the 4 year olds both processed more efficiently, as reflected by the positive scores, correctly written characters than Tibetan alphabets (e.g., “ཨ”).

## Contribution of Age and Reading Ability to Stroke Sensitivity

Across age groups and with basic reaction time statistically controlled, partial correlation was conducted among stroke sensitivity, age (expressed in months as a continuous variable), non-verbal IQ, and Chinese word reading. Stroke sensitivity was significantly correlated with age,  $r = 0.37$ ,  $p = 0.005$ , and Chinese word reading,  $r = 0.42$ ,  $p = 0.001$ , but not with non-verbal IQ,  $p = 0.169$ .

To better understand the relative contribution of age (maturation) and word reading ability (reading experience) to stroke sensitivity, hierarchical regression analyses were conducted with the dependent variable being stroke sensitivity and the independent predictors being age and Chinese word reading entered in different steps. We first examined whether the assumptions of regression were met, namely, linearity, normality, homoscedasticity, and independence. Visual inspection of the P-P plot confirmed the normality of error distribution. Examination of the scatter plot of residuals confirmed the linearity of relationship and inspection of the plot of residuals versus predicted values confirmed the constancy of variance of the errors (homoscedasticity). The Durbin-Watson statistic being 2.37 (with 1.4–2.6 considered being ideal) supported the statistical independence of the errors. We then ran the hierarchical regression models. In model 1, Chinese word reading was entered in the first step and age was entered in the second step. In model 2, age was entered in the first step and Chinese word reading was entered in the second step. **Table 3** shows the results of the final models. In model 1, after statistically controlling for Chinese word reading, age did not explain significant additional variance in stroke sensitivity,  $r^2$  change = 0.04,  $F$ -change (1,54) = 2.63,  $p = 0.11$ . In model 2, after having statistically controlled for age, Chinese word reading explained significant additional 10% of variance in stroke sensitivity,  $F$ -change (1,54) = 6.89,  $p = 0.01$ .

## DISCUSSION

The present study examined 57 4- and 5-year-old Chinese children's sensitivity to stroke, the smallest unit of writing that carries no phonological or semantic information in Chinese. The stroke-level sensitivity were assessed through an orthographic matching task in which children were asked to judge whether two items displayed on a Surface Pro were exactly the same or different. Importantly, we recorded and analyzed children's reaction time, which was more objective and reliable than oral reports. Also, we explored the association of age and reading experience to the emergence of stroke sensitivity.

We found that stroke-level sensitivity emerges from age 5. The 5-year-olds processed stroke-missing/redundant characters more slowly than that of correctly written characters. Also, their stroke sensitivity score was significantly higher than that of the 4-year-olds. Importantly, the 5-year-olds did not demonstrate sensitivity to foreign script (Tibetan alphabets), indicating that their sensitivity to stroke was not due to the unusual look of the stroke-missing/redundant characters. Children with stroke sensitivity demonstrated poorer performance, in this study, on the stroke-missing/redundant character task than those without stroke sensitivity, which was intended and reflected in terms of slower processing (as indexed by longer reaction time) rather than lower accuracy rate. In other words, the “poorer” performance of children with stroke sensitivity on the stroke-missing/redundant characters does not mean that they cannot distinguish correctly written characters and stroke-missing/redundant character; rather, it means that children who are sensitive to strokes within the character are more easily disturbed (than children who have not yet developed such sensitivity) when they process stroke-missing/redundant character and thus take longer reaction time. This is in line with previous findings from studies examining letter processing in children from alphabetic writing systems. Developmental work on letter processing has demonstrated that while children become faster at processing both letters and non-letters with age (Gibson et al., 1962; Reitsma, 1978), children demonstrate improved performance for letters compared to non-letters as young as 6 years of age (Miller and Wood, 1995; Burgund et al., 2006). These results are also consistent with findings from Zhao and Li (2014) in which the researchers used a lexical decision task with four types of stimuli (Chinese characters, stroke combinations, character-like line drawings with stroke features removed, and general line drawings) and asked Chinese 3- to 6-year-olds to judge whether the stimulus was a real character or not. They found that children's awareness of stroke developed very fast

**TABLE 3 |** Results of the final models predicting stroke sensitivity from age and reading ability.

Model	Predictors	$R^2$	$R^2$ change	$B$	$SE\ B$	$\beta$
1	Step 1: Chinese word reading	0.18	0.18**	0.02	0.01	0.34*
	Step 2: Age	0.22	0.04	0.35	0.22	0.21
2	Step 1: Age	0.12	0.12*	0.35	0.22	0.21
	Step 2: Chinese word reading	0.22	0.10*	0.02	0.01	0.34*

\*\* $p < 0.01$ , \* $p < 0.05$ .

during 4–5 years and reached peak at age 5. The present study supported findings from Zhao and Li (2014) but provided more objective, thus stronger, evidence for the emergence of stroke-level sensitivity from age 5 in Chinese children.

We also found that stroke sensitivity was linked more to reading experience than to maturation. Results of hierarchical regression analyses showed that Chinese word reading explained significant additional variance in stroke sensitivity after having statistically controlled for age, but age did not explain significant additional variance in stroke sensitivity beyond reading experience. This finding is consistent with previous studies showing that reading experience, rather than age, plays a more important role in learning to read among school-age children (e.g., Burgund et al., 2006; Zhao et al., 2014a). We expected that at an earlier stage of learning to read, e.g., in kindergarten years, maturation should play an equally important role, considering that perception of objects and events in the natural environment is a real-time task and a child needs to have mature sensory primitives (Aslin and Smith, 1988), but finding of the current study did not support this expectation. We found that although the young children's retinal processes is still developing, their visual acuity enables them to process fine features of written words (Gibson et al., 1962), e.g., stroke, the finest orthographic unit in the visually highly complex Chinese characters, and such sensitivity is significantly associated with their reading experience.

The current study sheds important light to the nature of early orthographic knowledge development. Similar to the alphabetic-language speaking counterparts (e.g., Lavine, 1977; Levy et al., 2006; Treiman et al., 2007; Pollo et al., 2009), Chinese kindergartners develop sensitivity to the identity and position of the components of writing they are exposed to before receiving formal literacy instruction. Different from previous studies that mostly examined components of writing that carry phonological or semantic information in varying degrees, the current study, for the first time to our knowledge, investigated beginning readers' sensitivity to the smallest unit of writing that carries no linguistic information at all in almost the most visually complex orthographies in the world. Our finding that children as young as age 5 can detect removal or addition of a single stroke from a character provides enlightening evidence that children develop formal knowledge about writing (visual-graphic and orthographic) independent of functional knowledge (phonological or semantic), and that visual processing of the holistic features of writing (which enables detection of a single change of stroke in a whole

character in the present study) is an important aspect of orthographic knowledge, especially in the early stage of learning to read.

Learning to read involves developing visual expertise for written form of words and linking this visual information to phonological and semantic information of words. It is the first step of visual-orthographic processing that is crucial for reading and learning to read (Maurer and McCandliss, 2007). Recent electrophysiological studies showed that visual expertise for written words correlates with children's individual reading ability (Zhao et al., 2014a) and that visual word expertise (written word N1) can be observed in young children who has not received formal reading training (Li et al., 2013). It was also reported that dyslexic children showed reduced N1 tuning for written words (Maurer et al., 2007). In future research, it is intriguing to explore early behavioral predictors, such as sensitivity to stroke, for later visual word expertise development. Such studies will help eventually elucidate the developmental and neural mechanisms underlying reading development.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Tsinghua University Research Ethics Committee. All subjects gave written informed consent from their parents.

## AUTHOR CONTRIBUTIONS

SL and LY collaboratively worked on the conception of the study, the acquisition, analysis, and interpretation of the data, and the writing-up of the paper.

## FUNDING

This research was funded by Open Research Fund of the Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences granted to LY and SL.

## ACKNOWLEDGMENT

We are grateful for the participating school and children.

## REFERENCES

- Anderson, R. C., Ku, Y., Li, W., Chen, X., Wu, X., and Shu, H. (2013). Learning to see the patterns in Chinese characters. *Sci. Stud. Read.* 17, 41–56. doi: 10.1080/1088438.2012.689789
- Aslin, R. N., and Smith, L. B. (1988). Perceptual development. *Annu. Rev. Psychol.* 39, 435–473. doi: 10.1146/annurev.ps.39.020188.002251
- Badian, N. (1994). Preschool prediction: orthographic and phonological skills, and reading. *Ann. Dyslexia* 44, 1–25. doi: 10.1007/BF02648153
- Burgund, E. D., Schlaggar, B. L., and Perterson, S. E. (2006). Development of letter-specific processing: the effect of reading ability. *Acta Psychol.* 122, 99–108. doi: 10.1016/j.actpsy.2005.11.005
- Cassar, M., and Treiman, R. (1997). The beginnings of orthographic knowledge: children's knowledge of double letters in words. *J. Educ. Psychol.* 89, 631–644. doi: 10.1037/0022-0663.89.4.631
- Chapman, L. J., and Chapman, J. P. (1973). Problems in the measurement of cognitive deficit. *Psychol. Bull.* 79, 380–385. doi: 10.1037/h0034541

- Chapman, L. J., Chapman, J. P., Curran, T. E., and Miller, M. B. (1994). Do children and the elderly show heightened semantic priming. How to answer the question. *Dev. Rev.* 14, 159–185. doi: 10.1006/drev.1994.1007
- CRT-Chinese version (1991). *Combined Raven's Test (CRT-Chinese Version)*. Shanghai: East China Normal University.
- Cunningham, A. E., Perry, K. E., and Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Read. Writ.* 14, 549–568. doi: 10.1023/A:1011100226798
- Demetriou, A., Kui, Z. X., Spanoudis, G., Christou, C., Kyriakides, L., and Platsidou, M. (2005). The architecture, dynamics, and development of mental processing: Greek, Chinese, or universal? *Intelligence* 33, 109–141. doi: 10.1016/j.intell.2004.10.003
- Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., and Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Curr. Biol.* 22, 814–819. doi: 10.1016/j.cub.2012.03.013
- Gibson, E. P., Gibson, J. J., Pick, A. D., and Osser, H. (1962). A developmental study of the discrimination of letter-like forms. *J. Comp. Physiol. Psychol.* 55, 897–906. doi: 10.1037/h0043190
- Goh, L., and Yap, V. B. (2009). Effects of normalization on quantitative traits in association test. *BMC Bioinformatics* 10:415. doi: 10.1186/1471-2105-10-415
- Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., and Fagot, J. (2012). Orthographic processing in baboons (*Papio papio*). *Science* 336, 245–248. doi: 10.1126/science.1218152
- Ho, C. S.-H., and Bryant, P. (1997). Learning to read Chinese beyond the logographic phase. *Read. Res. Q.* 32, 276–289. doi: 10.1598/RRQ.32.3.3
- Ho, C. S.-H., Yau, P. W.-Y., and Au, A. (2003). “Development of orthographic knowledge and its relationship with reading and spelling among Chinese kindergarten and primary school children,” in *Reading Development in Chinese Children*, eds C. McBride-Chang and H.-C. Chen (Westport, CT: Praeger), 47–67.
- Huang, H. S., and Hanley, J. R. (1994). Phonological awareness, visual skills and Chinese reading acquisition in first graders: a longitudinal study in Taiwan. *Adv. Study Chinese Lang. Process.* 1, 325–342.
- 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
- Lavine, L. O. (1977). Differentiation of letter like forms in prereading children. *Dev. Psychol.* 13, 89–94. doi: 10.1037/0012-1649.13.2.89
- Levy, B. A., Gong, Z. Y., Hessels, S., Evans, M. A., and Jared, D. (2006). Understanding print: early reading development and the contributions of home literacy experiences. *J. Exp. Child Psychol.* 93, 63–93. doi: 10.1016/j.jecp.2005.07.003
- Li, H., Peng, H., and Shu, H. (2006). A study on the emergence and development of Chinese orthographic awareness in preschool and school children. *Psychol. Dev. Educ.* 18, 35–38.
- Li, S., Lee, K., Zhao, J., Yang, Z., He, S., and Weng, X. (2013). Neural competition as a developmental process: early hemispheric specialization for word processing delays specialization for face processing. *Neuropsychologia* 51, 950–959. doi: 10.1016/j.neuropsychologia.2013.02.006
- Lonigan, C., Burgess, S., and Anthony, J. (2000). Development of emergent literacy and early reading skills in preschool children: evidence from a latent-variable longitudinal study. *Dev. Psychol.* 36, 596–613. doi: 10.1037/0012-1649.36.5.596
- Luo, Y. C., Chen, X., Deacon, S. H., and Li, H. (2011). Development of Chinese orthographic processing: a cross-cultural perspective. *Writ. Syst. Res.* 3, 69–86. doi: 10.1093/wsr/wsr008
- Luo, Y. C., Chen, X., Deacon, S. H., Zhang, J., and Yin, L. (2013). The role of visual processing in learning to read Chinese characters. *Sci. Stud. Read.* 17, 22–40. doi: 10.1080/10888438.2012.689790
- Maurer, U., Brem, S., Bucher, K., Kranz, F., Benz, R., Steinhausen, H.-C., et al. (2007). Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. *Brain* 130, 3200–3210. doi: 10.1093/brain/awm193
- Maurer, U., and McCandliss, B. D. (2007). “The development of visual expertise for words: the contribution of electrophysiology,” in *Single-word Reading: Biological and Behavioral Perspectives*, eds E. L. Grigorenko and A. J. Naples (Mahwah, NJ: Erlbaum), 43–64.
- McBride-Chang, C., Zhou, Y., Cho, J.-R., Aram, D., Levin, I., and Tolchinsky, L. (2011). Visual spatial skill: a consequence of learning to read? *J. Exp. Child Psychol.* 109, 256–262. doi: 10.1016/j.jecp.2010.12.003
- Miller, S. L., and Wood, F. B. (1995). Electrophysiological indicants of black-white discrimination performance for letter and non-letter patterns. *Int. J. Neurosci.* 80, 299–316. doi: 10.3109/00207459508986106
- Pollo, T. C., Kessler, B., and Treiman, R. (2009). Statistical patterns in children's early writing. *J. Exp. Child Psychol.* 104, 410–426. doi: 10.1016/j.jecp.2009.07.003
- Rasmussen, J. L., and Dunlap, W. P. (1991). Dealing with non normal data: parametric analysis of transformed data vs nonparametric analysis. *Educ. Psychol. Meas.* 51, 809–820. doi: 10.1177/001316449105100402
- Reitsma, P. (1978). Changes in letter processing in beginning readers. *J. Exp. Child Psychol.* 25, 315–325. doi: 10.1016/0022-0965(78)90085-1
- 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
- Siok, W. T., and Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Dev. Psychol.* 7, 886–899. doi: 10.1037/0012-1649.37.6.886
- Su, M., Wang, J., Maurer, U., Zhang, Y., Li, J., McBride, C., et al. (2015). Gene-environment interaction on neural mechanisms of orthographic processing in Chinese children. *J. Neurolinguistics* 33, 172–186. doi: 10.1016/j.jneuroling.2014.09.007
- Treiman, R., Cohen, J., Mulqueeny, K., Kessler, B., and Schechtman, S. (2007). Young children's knowledge about printed words. *Child Dev.* 78, 1458–1471. doi: 10.1111/j.1467-8624.2007.01077.x
- Valdois, S., Bosse, M. L., and Tainturier, M. J. (2004). The cognitive deficits responsible for developmental dyslexia: review of evidence for selective visual attentional disorder. *Dyslexia* 10, 339–363. doi: 10.1002/dys.284
- Van der Leij, A., van Bergen, E., van Zuijen, T., de Jong, P., Maurits, N., and Maassen, B. (2013). Precursors of developmental dyslexia: an overview of the longitudinal Dutch dyslexia programme study. *Dyslexia* 19, 191–213. doi: 10.1002/dys.1463
- Wang, Y., Yin, L., and McBride, C. (2015). Unique predictors of early reading and writing: a one-year longitudinal study of Chinese kindergartners. *Early Child. Res. Q.* 32, 51–59. doi: 10.1016/j.jecresq.2015.02.004
- Yin, L., and McBride, C. (2015). Chinese kindergartners learn to read characters analytically. *Psychol. Sci.* 26, 424–432. doi: 10.1177/0956797614567203
- Zhao, J., Kipp, K., Gaspar, C., Maurer, U., Weng, X., Mecklinger, A., et al. (2014a). Fine neural tuning for orthographic properties of words emerges early in children reading alphabetic script. *J. Cogn. Neurosci.* 26, 2431–2442. doi: 10.1162/jocn\_a\_00660
- Zhao, J., and Li, S. (2014). Development of form recognition of visual word in Chinese Children of 3 to 6 Years Old. *Psychol. Sci.* 37, 357–362.
- Zhao, J., Zhao, P., Weng, X., and Li, S. (2014b). Do preschool children learn to read words from environmental prints? *PLoS ONE* 9:e85745. doi: 10.1371/journal.pone.0085745
- Zhou, Y., McBride-Chang, C., and Wong, N. (2014). What is the role of visual skills in learning to read? *Front. Psychol.* 5:776. doi: 10.3389/fpsyg.2014.00776

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Li and Yin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Applicability of the Compensatory Encoding Model in Foreign Language Reading: An Investigation with Chinese College English Language Learners

**Feifei Han\***

*Centre for Research on Learning and Innovation, Sydney School of Education and Social Work, The University of Sydney, Sydney, NSW, Australia*

## OPEN ACCESS

### **Edited by:**

Giseli Donadon Germano,  
Universidade Estadual Paulista, Brazil

### **Reviewed by:**

Yuejin Xu,  
Murray State University, USA  
Mercedes Inda-Caro,  
University of Oviedo, Spain

### **\*Correspondence:**

Feifei Han  
feifei.han@sydney.edu.au

### **Specialty section:**

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 30 January 2017

**Accepted:** 19 April 2017

**Published:** 04 May 2017

### **Citation:**

Han F (2017) Applicability of the  
Compensatory Encoding Model  
in Foreign Language Reading: An  
Investigation with Chinese College  
English Language Learners.  
Front. Psychol. 8:681.  
doi: 10.3389/fpsyg.2017.00681

While some first language (L1) reading models suggest that inefficient word recognition and small working memory tend to inhibit higher-level comprehension processes; the Compensatory Encoding Model maintains that slow word recognition and small working memory do not normally hinder reading comprehension, as readers are able to operate metacognitive strategies to compensate for inefficient word recognition and working memory limitation as long as readers process a reading task without time constraint. Although empirical evidence is accumulated for support of the Compensatory Encoding Model in L1 reading, there is lack of research for testing of the Compensatory Encoding Model in foreign language (FL) reading. This research empirically tested the Compensatory Encoding Model in English reading among Chinese college English language learners (ELLs). Two studies were conducted. Study one focused on testing whether reading condition varying time affects the relationship between word recognition, working memory, and reading comprehension. Students were tested on a computerized English word recognition test, a computerized Operation Span task, and reading comprehension in time constraint and non-time constraint reading. The correlation and regression analyses showed that the strength of association was much stronger between word recognition, working memory, and reading comprehension in time constraint than that in non-time constraint reading condition. Study two examined whether FL readers were able to operate metacognitive reading strategies as a compensatory way of reading comprehension for inefficient word recognition and working memory limitation in non-time constraint reading. The participants were tested on the same computerized English word recognition test and Operation Span test. They were required to think aloud while reading and to complete the comprehension questions. The think-aloud protocols were coded for concurrent use of reading strategies, classified into language-oriented strategies, content-oriented strategies, re-reading, pausing, and meta-comment. The correlation analyses showed that while word recognition and working memory were only significantly related to



frequency of language-oriented strategies, re-reading, and pausing, but not with reading comprehension. Jointly viewed, the results of the two studies, complimenting each other, supported the applicability of the Compensatory Encoding Model in FL reading with Chinese college ELLs.

**Keywords:** Compensatory Encoding Model, word recognition, working memory, reading comprehension, foreign language reading, Chinese college English language learners

## INTRODUCTION

Most of us read everyday, from academic texts to technical reports, from literature to popular magazines, and from newspapers to brochures. The seemingly common practice of reading in fact is a highly complex cognitive activity (Koda, 2007; Yamashita, 2013), even when reading in one's first language (L1), let alone reading in a foreign language (FL). One of the major obstacles faced by FL readers is slow word recognition, which requires much conscious deliberation (Segalowitz, 2000, 2003; Fukkink et al., 2005; Segalowitz and Hulstijn, 2005). Inefficient word recognition takes much cognitive resources, such as working memory, which is essential for reading comprehension to occur (Juffs and Harrington, 2011). Some theoretical models of reading, such as the Verbal Efficiency Model highlights the importance of word recognition efficiency and working memory, suggesting that inefficiency in word recognition and small working memory tend to inhibit higher-level comprehension processes (e.g., Perfetti, 1985, 2007). On the other hand, in the Compensatory Encoding Model, the role of strategic processing is emphasized and the model postulates that as long as readers have sufficient time to carry out a reading task, slow word recognition and limited working memory do not normally hinder reading comprehension, because readers are able to apply some kinds of higher-order metacognitive strategies to remedy processing efficiency (slow word recognition) and resource limitation (small working memory), and that is to say those metacognitive strategies have compensatory characteristics (Walczyk, 2000; Walczyk et al., 2001, 2007).

The Compensatory Encoding Model of reading was proposed by Walczyk and his colleagues to explicate "the interplay between automatic and control processes" (Walczyk, 2000, p. 35) in L1 reading beyond the initial stages of learning to read (Walczyk, 1993, 1995; Walczyk and Taylor, 1996; Walczyk et al., 2001, 2007). The construction of the model is based on a number of L1 reading theories (Walczyk, 2000), including Automaticity Theory (LaBerge and Samuels, 1974), the Verbal Efficiency Model (Perfetti, 1985, 1988), Metacognitive Theory (Baker and Brown, 1984), Constructively Responsive Theory (Pressley and Afflerbach, 1995), and Rauding Theory (Carver, 1997). According to the model, in fluent reading, processes such as identifying words and accessing to meanings tend to be carried out automatically, which make few demands on working memory. As a result, working memory can be freed up for higher-level comprehension processes, which are operated in slow, error prone, unstable, and serial manner (Walczyk, 2000; Shiotsu, 2009). In situations where readers

have processing limitation (i.e., inefficient word recognition) and have resource limitation (i.e., small working memory), the model highlights the importance of compensatory mechanisms, which are metacognitive strategic processing (Walczyk, 1995). The model postulates that the condition for compensatory mechanisms to operate successfully during reading relies heavily on reading time: when there is severe time constraint, such as in a testing situation, it is less likely for readers to operate compensatory mechanisms freely. The application of mechanism as successful compensation for inefficient word recognition and small working memory tends to occur when reading without much time constraint (Walczyk, 1993, 1995, 2000; Walczyk and Taylor, 1996; Walczyk et al., 2001, 2007).

The Compensatory Encoding Model entails two important predictions. The first prediction is that reading conditions varying time influence the relationship between word recognition efficiency, working memory, and reading comprehension. When reading without time constraint, inefficient word recognition and small working memory do not normally affect reading comprehension "because compensatory mechanisms operate routinely during performance" (Walczyk, 1993, p. 127). That is to say there will be only a weak or no relationship between word recognition efficiency, working memory, and comprehension. When reading is under time constraint, compensatory mechanisms are less likely to operate freely; hence, inefficient word recognition and small working memory tend to be adversely affects reading comprehension (Walczyk, 1993, 1995, 2000; Walczyk and Taylor, 1996; Walczyk et al., 2001, 2007).

The second prediction is that when reading occurs without time constraint, it is the use of compensatory strategies which are predictive of reading comprehension. This means that readers with slower word recognition and smaller working memory tend to use more metacognitive mechanism in reading without time constraint, and as a result, their reading comprehension tends not to be affected by word recognition inefficiency and limited working memory (Walczyk and Taylor, 1996).

To empirically test the Compensatory Encoding Model, Walczyk and his associates conducted a series of studies with both young and mature native English speakers. In one of the earlier studies, Walczyk (1995) compared contributions made by word recognition and working memory to reading comprehension with and without time constraint among university students. The results showed that without time constraint, none of the measures of word recognition and working memory related to reading comprehension. However, when reading under time constraint, word recognition efficiency and working memory were significantly associated with comprehension. This study

provided evidence for the first prediction in the Compensatory Encoding Model among adult L1 readers.

To test the actual use of compensatory strategies in reading (the second prediction), Walczyk and his colleagues conducted a few studies with children and adults. Some of these studies provided full support (e.g., Walczyk et al., 2004), whereas some studies only provided partial support for the Compensatory Encoding Model, especially among younger children (e.g., Walczyk et al., 2007). Among primary school students, Walczyk et al. (2004) recorded the participants' reading aloud of one narrative and one expository text by giving children sufficient time to read. They found that measures of word recognition efficiency and working memory were significantly and negatively related to frequency of using pausing and re-reading – two kinds of compensatory strategies, in reading both types of texts. Walczyk and his associates also found some support for the Compensatory Encoding Model among adult readers (e.g., Walczyk and Taylor, 1996; Walczyk et al., 2001). For instance, Walczyk and Taylor (1996) asked university students to read texts on a computer screen without time constraint and recorded students' re-reading behaviors using a computer program. Their re-reading was found to be significantly correlated with the speed measures of word recognition and working memory, suggesting that readers with inefficient word recognition and small working memory tended to re-read more frequently.

Using the think-aloud method, which allowed readers to process reading task at hand with ample time, Walczyk et al. (2001) found that slower word recognition was associated with more frequent pausing, looking back, and re-reading behaviors, and slower speed measure of working memory was associated with more re-reading behaviors. In addition, neither of the speed measure of word recognition nor working memory was related to reading comprehension, implying readers' compensatory use of pausing, looking back, and re-reading strategies in reading.

A further examination of the Compensatory Encoding Model investigated developmental pattern in the relationship between word recognition, working memory in relation to use of compensatory strategy among third, fifth, and seventh graders (Walczyk et al., 2007). The researchers manipulated the reading conditions by placing time restrictions to create either time restricted reading or non-time restricted reading and students were randomly assigned to one of the reading conditions. In the non-time restricted reading, students were asked to read-aloud to enable coding of possible compensatory strategies, such as pausing, looking back, and jumping over. The results demonstrated that the relational pattern between word recognition, working memory, and use of compensatory reading strategies exhibited a consistent pattern across the three grades: both the accuracy measure of word recognition and working memory were negatively correlated with jumping over for third and seventh graders, and with looking back for fifth graders, indicating that slower word recognition and smaller working memory appeared to be associated with more frequent application of compensatory mechanism in reading.

However, the relationship between word recognition, working memory, and reading comprehension in the non-time restricted reading condition displayed different patterns for different

grades. While word recognition and working memory was found to adversely affect reading comprehension for the third and fifth graders, they did not affect reading comprehension for the seventh graders. The lack of relation for seventh graders seemed to indicate that strategy use as successful compensation was only realized among older and experienced readers, who were more metacognitively and strategically oriented than younger and less experienced readers. The developmental pattern of the results are in line with the creation of the model, which targets more experienced readers, who are able to orchestrate metacognitive reading strategies strategically.

In summary, there was ample empirical evidence which supported the two predictions of the Compensatory Encoding Model in L1 reading: (1) reading conditions varying time affect the relation between word recognition, working memory, and comprehension; (2) in non-time constraint reading, experienced L1 readers displayed compensatory nature of reading strategy use for word recognition inefficiency and working memory limitation.

In FL reading, testing of the Compensatory Encoding Model is lacking. To the best of our knowledge, the only study which has directly investigated the Compensatory Encoding Model in FL reading is conducted by Stevenson (2005) with 22 Dutch adolescent English language learners (ELLs). Stevenson (2005) measured word recognition speed using a computerized lexical decision task, and adopted a think-aloud method to measure concurrent reading strategy use. Reading while thinking-aloud gave readers sufficient time to process a text, which simulated non-time constraint reading condition for the use of reading strategies as a remedy of word recognition inefficiency and limited working memory capacity in the Compensatory Encoding Model. The coding of reading strategy use were broken down into three dimensions, namely orientation of processing (language and content), type of processing (metacognitive, cognitive, and cognitive-iterative), and domain of processing (above clause, clause, and below clause). The results showed that word recognition speed correlated with language-oriented strategies (i.e., strategies directed toward linguistic information), cognitive strategies (i.e., strategies involving direct mental processing of a text), and clause and above-clause level strategies (i.e., strategies which help understand whole or a few successive clauses), but not with reading comprehension. These results provided empirical evidence that in the non-time constraint reading, similar to L1 readers, FL readers are able to deploy compensatory use of reading strategies for slow word recognition so that word recognition did not adversely affect reading comprehension and frequency of strategy use which are characterized as compensatory nature enabled them to achieve reasonable comprehension.

A number of issues in Stevenson's (2005) study warrant further testing of the Compensatory Encoding Model in FL reading. First of all, while Stevenson's study focused on investigating the second prediction in the Compensatory Encoding Model, it did not directly examine the first prediction, that is reading conditions varying time affects the relation between word recognition, working memory, and comprehension in FL reading. Secondly, working memory was not examined in the study, so it was not

able to draw any direct conclusion regarding the relationship between FL readers' resource limitations, strategy use, and reading comprehension. Thirdly, the word recognition measure in Stevenson's study was not appropriate for FL readers, as it only involved decoding but meaning access. For FL readers, decoding may activate a connection to meaning or only leads to a weak connection to meaning (Shaw and McMillion, 2008; Grabe, 2009). Thus, it is more appropriate to use a task which requires meaning access through identification of word forms to measure word recognition efficiency among FL readers. In addition, Dutch ELLs speak a L1 which is typologically close to English. The similarity between the two languages may pose little difficulty in English word recognition. Considering ample evidence of qualitatively different cognitive processes for word recognition by alphabetic and non-alphabetic learners (Koda, 1994, 1996, 2005, 2007), it is necessary to examine whether FL readers whose L1 is a non-alphabetic language, such as Chinese, are able to deploy compensatory use of reading strategy for word recognition inefficiency and working memory limitation as those native English speakers and those ELLs whose L1 is also an alphabetic language.

The present research aims to test the applicability of the Compensatory Encoding Model in FL reading with Chinese ELLs. Two studies were conducted, each focusing on testing one of the important predictions in the Model. Study one aimed to test whether reading condition varying time affects the relationship between word recognition efficiency, working memory, and reading comprehension. The research questions for study one are: (1) To what extent does word recognition efficiency and working memory relate to reading comprehension in the time constraint and non-time constraint FL reading? (2) To what extent does word recognition efficiency and working memory contribute to reading comprehension in the time constraint and non-time constraint FL reading? According to the Compensatory Encoding Model, we hypothesized that in FL reading, when the reading condition is strictly time limited, both word recognition efficiency and working memory tend relate to and contribute to reading comprehension. When reading condition allows readers to have ample time to complete a reading task, word recognition efficiency and working memory tend not to associate with and contribute to reading comprehension.

Study two aimed to examine whether Chinese ELLs are able to operate metacognitive reading strategies as a compensatory way of reading comprehension for inefficient word recognition and working memory limitation in non-time constraint reading. The research question for study two is: What is the interrelationship between word recognition efficiency, working memory, use of reading strategies, and reading comprehension in non-time constraint FL reading?

## STUDY ONE

### Material and Methods

#### Participants

The participants in study one were 402 second year undergraduates (138 males and 266 females) recruited from

a national university in China. We targeted second year students because first year students had just entered the university, and students beyond second year are not required to be enrolled in the compulsory college English learning (Hu, 2005). The recruitment focused on non-English major undergraduates, as English majors are not representative of the majority of Chinese ELLs, due to their presumably better proficiency and greater interest in English learning. The participants came from 12 English classes, majored in eight disciplines (i.e., Economics and Business, Humanities and Social Sciences, Information Technology and Computer Science, Material Engineering, Mechanical Engineering, Science, Printing and Packaging Technology, and Water Resources and Hydraulic Power). Their age ranged between 18 and 22 with a Mean (*M*) of 20.22 and a Standard Deviation (*SD*) of 0.93. On average, the participants received 7.5 years of English instruction.

### Materials

#### Word recognition test and scoring

To measure word recognition efficiency, we used a computerized test, which required learners to decide as quickly as possible whether a pair of words had a similar meaning (synonyms) or had opposite meanings (antonyms). The format of the test was adapted from Haynes and Carr's (1990) paper-and-pencil test and it was essential for the participants to access the meaning of words using this test to measure word recognition. The testing items were 60 word pairs of four different parts of speech (i.e., noun: 14, verb: 20, adjective: 20, and adverbs: 6). To avoid testing vocabulary knowledge of the ELLs, all the testing items were from the most frequent 2,000 words band in the British National Corpus word list (Nation, 2004). The lexical relationship between the words (synonyms or antonyms) was checked using an online thesaurus<sup>1</sup>.

The test which was delivered through Lenovo computers with 17 inch screen with Windows XP system using DMDX software (version 3.3.1.1) (Forster and Forster, 2003) was held in a quiet computer laboratory. Students were required to make a judgment as quickly as possible by pressing two keys marked with Synonym or Antonym. The order of the testing items was randomized using the random function of the software. The test instructions were given in Chinese, and the test started with six practice pairs (see Appendix 1 for sample items).

As the word recognition test aimed to measure students' efficiency of recognizing English words rather than to test their English vocabulary knowledge, only values of reaction time rather than correctness of judgment were used for data analysis. As in most reaction time analyses, the upper and lower thresholds were set as 3 *SDs* above and below the *M* reaction time of each item (e.g., Muljani et al., 1998; Koda, 2000; van Gelderen et al., 2004). The values of reaction time falling outside the thresholds were located and transformed into missing values. The missing values (accounting for 0.94%) were estimated by using the Expectation Maximization algorithm. After the estimation, the Cronbach's alpha was calculated and its value was 0.94, indicating good reliability.

<sup>1</sup>www.thesaurus.com

### Working memory test and scoring

We used a modified computerized Operation Span Task, which was developed by Unsworth et al. (2005). The reasons for choosing the OST over the popular Reading Span Task (RST) were: (1) the OST does not involve testing the participants' reading comprehension ability whereas the RST requires readers to process sentences for comprehension (Kintsch, 1998; Seigneuric et al., 2000; Koda, 2005; Alptekin and Erçetin, 2010, 2011; Rai et al., 2011); (2) the OST tends not to be influenced by language proficiency (Service et al., 2002), whereas the RST if tested in a FL, is affected by levels of proficiency in that FL; and (3) the OST, which is a well-established measure of working memory in the field of psychology with confirmed validity and reliability (Conway et al., 2005; Unsworth et al., 2005), examines both storing and processing functions of working memory simultaneously (Baddeley, 2006, 2007; Juffs and Harrington, 2011).

The OST asked participants to memorize isolated English words displayed on the computer screen for half a second at the same time to judge the correctness of a simple mathematical equation involving addition, subtraction, multiplication, and/or division (e.g.,  $(10 \times 5) - 20 = 30$ ) by pressing a key marked with Correct or Wrong. The test was also delivered using the DMDX software (version 3.3.1.1) (Forster and Forster, 2003) via the same computers as used for the word recognition test. There were 40 items (i.e., one item consists of a word for recall and a mathematical equation for judgment) divided into 10 sets ranging from 2 to 6 items in each set. After each set, when the computer displayed "Recall and write down the words within the set (in Chinese)," the participants were asked to write the words on an answer sheet. Upon completion of writing, they needed to press the Space key to proceed to the next set. The words for recall (20 nouns and 20 verbs) were from the most frequent 300 words in British National Corpus word list (Nation, 2004), and all the words were only one syllable ranging from 4 to 6 letters. The task instructions were given in Chinese, and the test started with 14 practice items divided into 4 sets (see Appendix 2 for sample items).

To score the working memory test, we used composite Z-scores, which were formed by averaging Z-scores of: (1) the number of correctly recalled words, (2) the number of correct judgment, and (3) the reaction time of the judgment (Waters and Caplan, 1996). For the reaction time of the judgment, we trimmed the data using the thresholds of 3 SDs above and below the *M* reaction time of each item. The outlying values of reaction time were marked as missing values (accounting for about 1%), and were estimated by the Expectation Maximization algorithm. We multiplied the reaction time by  $-1$  in order for a higher value to represent better performance and then transformed them into Z-scores (Leeser, 2007). The reliability, which was calculated using the composite Z-scores, had a value of 0.85, indicating that the working memory test was reliable.

### Reading comprehension tests and scoring

Two parallel reading comprehension tests in the two reading conditions (i.e., time-constraint vs. non-time constraint) were

customarily compiled using four expository texts. As research has shown that text type influences reading strategy use (Alderson, 2000; Horiba, 2000; Grabe, 2009; Alptekin and Erçetin, 2011), a decision was made to use a single text type to avoid text type being a confounding factor. Expositions were chosen because they were most familiar to the participants according to the participants' English teachers. The four texts were adapted from *College Reading Workshop* (Malarcher, 2005), a reading practice book targeting upper-intermediate learners of English. We chose texts based on the following criteria by consulting the participants' English teachers: (1) the linguistic difficulty was not overly challenging in terms of lexical and morphosyntactic complexity; (2) understanding the texts did not require specialized background knowledge; and (3) the texts were interesting for the participants to read.

To ensure that the two tests had a similar level of readability, the following efforts were made. (1) The topics of the texts in the two reading conditions were matched: both tests had one text related to human body [Text 1 (T1): *Fat to Store or Fat to Burn?*, Text 3 (T3): *Ideas about Beauty*]; and the other related to technology [Text 2 (T2): *Commerce through the Internet*, Text 4 (T4): *A Second Look at Virtual Advertisements*]. (2) Each text had six paragraphs and similar word counts (T1 to T4: 588, 594, 588, and 596 words). (3) The four texts were comparable in terms of T-units and average number of words per T-unit (T1: 29 T-units, 20.24 words/T-unit; T2: 29 T-units, 20.38 words/T-unit; T3: 29 T-units, 20.34 words/T-unit; and T4: 29 T-units, 20.00 words/T-unit). (4) The four texts had comparable Flesch-Kincaid Grade level (T1 to T4: 9.40, 10.90, 9.80, and 10.90). The Flesch-Kincaid Grade level is a commonly used readability index for native English speakers rather than ELLs, our purpose of using it was only to check whether the four texts had similar levels of readability.

Reading comprehension questions were multiple choice, which is the most commonly used approach for assessing reading comprehension (Kendall et al., 2001; Brantmeier, 2005; Alptekin, 2006; Phakiti, 2007; Alptekin and Erçetin, 2011). Using multiple choice has a number of advantages over other methods (Brantmeier, 2005; Phakiti, 2007): (1) it is objective and does not require training for scoring; (2) it is convenient to administer to large numbers of students; (3) it does not make heavy demands on readers' memory, writing ability, and synthesizing skills, compared with recall and cloze tasks (Jenkins et al., 2004; Johnson et al., 2005; Alptekin, 2006; Alptekin and Erçetin, 2011). Although multiple choice question has been criticized for guessing (Alderson, 2000), we informed students that the tests were not part their assessment for the course so that they should avoid guessing.

For each text, 10 multiple choice questions with four possible choices were constructed. Five of them were literal comprehension questions requiring specific information directly stated in the text; and five were inferential questions measuring global text comprehension, such as theme and aims of the text. A correct answer for one question received one point, and the maximum achievable score for each test was 20 (see the Appendix 3 for a sample text and comprehension questions).



## Data Collection Procedure

The participants were allowed to spend 40 min to complete the reading comprehension test in the time constraint reading condition. The time was decided by allowing 60% of the average normal reading time of 100 students in a pilot study following Walczyk (1995) and Walczyk and Taylor (1996)'s suggestion. The students in the pilot study had similar level of reading proficiency as the participants. The two reading comprehension tests were conducted on 1 day in the students' classrooms. Students completed the word recognition and working memory tests on a separate day in a computer laboratory. The 402 participants were grouped into 20 groups with approximately 20 students in each group. Students in each group started the tests at the same time but the completion time varied among them. On average, they spent 30 min to complete the two tests.

## Data Analysis

To answer research question 1 – the relationship between word recognition, working memory, and reading comprehension in the time constraint and non-time constraint reading conditions, correlation analyses were applied separately for the two reading conditions. To answer research question 2 – the contribution of word recognition and working memory to reading comprehension in the time constraint and non-time constraint reading conditions, regression analyses were performed.

## Results

### Results of Research Question 1

**Table 1** presents the descriptive statistics of all the tests in study one. The descriptive statistics of the word recognition test are reaction time in milliseconds, descriptive statistics of the working memory test are composite *Z*-scores comprised of accuracy of judgment, the reaction time of judgment, and the number of correctly recalled words.

**Table 1** shows that the *M* scores of the non-time constraint reading comprehension were higher than that in the time constraint reading. A one-way repeated ANOVA revealed that the difference was significant,  $F(1,401) = 127.48, p < 0.01$ , indicating that when FL readers read without time constraint, they achieved better comprehension, even though the texts in the two reading conditions were matched for the level of readability and topics.

The results of the correlation analyses are displayed in **Table 2**, which shows a small and negative correlation between word recognition and reading comprehension in the time constraint condition ( $r = -0.22, p < 0.01$ ), indicating that the slower one's word recognition (hence longer reaction time) was, the poorer one's comprehension was in the time constraint reading. However, we found that the correlation between word recognition and reading comprehension in the non-time constraint condition was not significant ( $r = -0.09, p = 0.07$ ), suggesting that when the participants were allowed to read with sufficient time, the speed of recognizing English words did not affect their reading comprehension.

**Table 2** further shows that working memory was significantly and positively correlated with reading comprehension in both time constraint ( $r = 0.20, p < 0.01$ ) and non-time constraint

reading ( $r = 0.11, p < 0.05$ ). To compare the strength of the correlations between working memory and reading comprehension in the two reading conditions, we used Steiger's (1980) *Z*-test, and the results indicated that the strength of the correlation in the time constraint reading was significantly stronger than that in the non-time constraint reading,  $Z = 1.71, p < 0.05$ . The results indicated that readers with larger working memory capacity tended to be more strongly related to better reading comprehension in time constraint reading than in non-time constraint reading.

### Results of Research Question 2

As the word recognition did not correlate with reading comprehension in the non-time constraint condition, the reading comprehension was regressed on working memory only and the results are presented in **Table 3**. **Table 3** revealed that even though working memory significantly contributed to the regression model:  $F(1,400) = 2.26, p < 0.05, f^2 = 0.01$ , it accounted for only 1% of variance in reading comprehension in non-time constraint reading with negligible effect size, suggesting the rather weak predictive power of working memory capacity to comprehension in the non-time constraint reading.

For the time constraint condition, the reading comprehension was regressed on word recognition and working memory. Before performing the multiple regression analysis, we conducted a series of tests to examine the essential assumptions for reliable results. First, the correlation between the two predictors (word recognition and working memory) was 0.28, and the values of Tolerance for word recognition = 0.92, and working memory = 0.92; these ensured no multicollinearity. Second, the analysis of standard residuals on the data identified four cases as outliers (falling out of  $\pm 3$  SDs), which were removed (less than 1% of the data). Third, the value of the Durbin-Watson was 1.86, suggesting that there was no auto-correlation in our data. After confirmed that our data met these assumptions, we proceeded with the multiple analysis. The results were presented in **Table 4**, which shows that both word recognition ( $\beta = -0.18, R^2 = 0.05, p < 0.01, f^2 = 0.05$ ) and working memory ( $\beta = 0.15, R^2 = 0.02, p < 0.01, f^2 = 0.02$ ) significantly contributed to reading comprehension in the time constraint reading, explaining 5 and 2% of variance respectively.

In summary, we found that word recognition and working memory related to and contributed to reading comprehension in the time constraint and non-time constraint reading conditions differently among Chinese college ELLs. The results indicated that varying reading time affected the relationship between word recognition, working memory, and reading comprehension. This is consistent with L1 findings and provided some empirical evidence for the first prediction of the Compensatory Encoding Model among FL readers.

## Discussion

The results of study one demonstrate that the correlations between word recognition, working memory, and reading comprehension in the time constraint and non-time constraint reading conditions differ from each other. The results were consistent with Walczyk's (1995) study with L1 readers and

**TABLE 1 | Descriptive statistics of all the tests in study one.**

Variables	<i>M</i>	<i>SD</i>	Minimum	Maximum
Word recognition ( <i>in milliseconds</i> )	1981.06	463.59	1026.95	3984.58
Working memory	0.00	0.19	−0.78	0.40
Comprehension in time constraint condition	12.99	2.88	2.00	19.00
Comprehension in non-time constraint condition	15.14	2.42	2.00	20.00

**TABLE 2 | Results of correlation analyses.**

Variables	Working memory	Comprehension in non-time constraint reading	Comprehension in time constraint reading
Word recognition	−0.28**	−0.09	−0.22**
Working memory	—	0.11*	0.20**

\*\* $p < 0.01$ , \* $p < 0.05$ , (*two-tailed*).

**TABLE 3 | Results of simple regression analysis.**

Variables	<i>B</i>	$\beta$	<i>t</i>	<i>p</i>	<i>r</i> <sup>2</sup>
Constant	0.76	—	125.14	0.00	—
Working memory	0.07	0.11*	2.26	0.02	0.01

$R^2 = 0.01$ , \* $p < 0.05$ .

**TABLE 4 | Results of multiple regression analysis.**

Variables	<i>B</i>	$\beta$	<i>t</i>	<i>p</i>	<i>r</i> <sup>2</sup>
Constant	0.67	—	89.05	0.00	—
Word recognition	−0.10	−0.18**	−4.53	0.00	0.05
Working memory	0.11	0.15**	2.92	0.00	0.02

$R^2 = 0.05$ , \*\* $p < 0.01$ , for word recognition; and  $R^2 = 0.02$ , \*\* $p < 0.01$ , for working memory.

extended the first prediction in the Compensatory Encoding Model that varying reading time affects relationship between word recognition efficiency, working memory capacity, and reading comprehension to FL reading. As indicated in the Compensatory Encoding Model, reading varying time may determine whether readers are able to orchestrate strategic processing successfully as a means to compensate for inefficient word recognition and working memory limitation. This seemed to be the case in FL reading with readers whose L1 and FL have different orthography.

No studies in FL reading have compared the time constraint and non-time constraint reading when they have examined the relationship between word recognition, working memory, and reading comprehension. However, research in FL reading on the relationship between word recognition and reading comprehension has produced inconsistent results: while some studies revealed that individual differences in word recognition did not lead to different levels of text comprehension in FL reading (e.g., Haynes and Carr, 1990; Shiotsu, 2009; Yamashita, 2013); other studies found that word recognition inefficiency adversely affected comprehension in FL reading (e.g., Nassaji and Geva, 1999; Nassaji, 2003a; Tsai, 2008). These inconclusive

findings may be attributed to lack of control of reading time in these studies. The studies which found an inhibitory effect of inefficient word recognition might have readers complete reading tasks within restricted reading time in their design. For instance, Nassaji (2003a) and Nassaji and Geva (1999) reported a significant relationship between word recognition and reading comprehension for ELLs with advanced reading proficiency. The reading comprehension test used in the two studies is the Nelson–Denny Reading Test, a standardized L1 reading test for native English speakers. Thus, the test might be challenging for the ELLs, hence reading the texts might require additional time than that had been given to their participants.

On the other hand, those studies which reported no relationship between word recognition and reading comprehension might have given readers ample time to complete reading tasks. For example, Haynes and Carr (1990) found that individual differences in word recognition did not lead to different levels of text comprehension in FL reading among Chinese ELLs. In this study, the participants were asked to read a 500-word text and to record their reading time, which in fact, allowed students read at their own pace. Lack of control for time in reading might simulate the non-time constraint reading, which is essential for readers to utilize strategic repertoire as compensation for slow word recognition. Similarly, among Japanese ELLs, Yamashita (2013) also allowed readers to read in their own pace and recorded readers' reading rates, and she found that word recognition was only significantly related to reading rate but not with reading comprehension.

Our results that varying time affects the relationship between word recognition efficiency and reading comprehension may also offer some explanations as to why some intervention studies on word recognition training failed to show any effect of increased word recognition on reading comprehension (e.g., Fukkink et al., 2005; Akamatsu, 2008). Even though the participants in the Fukkink et al. (2005) improved word recognition efficiency after intervention, their reading comprehension did not show any improvement. As explained by the researchers that “t[T]here was no time limit for completing the reading test” (Fukkink et al., 2005, p. 64), and we speculate this could be the reason for failure of showing the effect of increased word recognition efficiency on text comprehension. In the future, researchers may consider controlling reading time when testing comprehension in order to more accurately reflect the effect of word recognition training.

We found that even in the time constraint reading condition, word recognition efficiency and working memory could only

make small contributions to reading comprehension. As reading is a highly complex cognitive activity and comprehension depends substantially on other skills in addition to word recognition and working memory, such as L1 reading ability, metacognitive knowledge, and FL proficiency (e.g., grammatical and vocabulary knowledge) (van Gelderen et al., 2003, 2004; Koda, 2005; van Gelderen et al., 2007; Grabe, 2009; Grabe and Stoller, 2011). Previous studies demonstrated that linguistic proficiency and metacognitive knowledge make much more substantial contribution than word recognition and sentence processing speed to FL reading (e.g., van Gelderen et al., 2003, 2004, 2007). Thus, it was reasonable to expect weak predictive power of word recognition and working memory.

In terms of the relationship between working memory and FL reading comprehension, our research found that students who have a larger working memory were more likely to achieve better comprehension in FL reading. The significant relationship in our research is in line with previous FL reading studies, such as Harrington and Sawyer's (1992), and Walter's (2004) studies, though the strength of the relationship in our research is much weaker than theirs (Harrington and Sawyer:  $r = 0.57$ ; Walter:  $r = 0.79$  for lower-intermediate group;  $r = 0.46$  for higher-intermediate group). One possible reason for the difference could be the different instruments used for measuring working memory. Both Harrington and Sawyer's and Walter's used the RST to measure working memory, which largely depends on one's reading ability, and the working memory tends to be confounded with reading comprehension ability (Seigneuric et al., 2000; Koda, 2005), hence may be responsible for much stronger association. Our research used the OST to measure working memory, which reduced the involvement of reading ability to its minimum. Therefore, the significant and small relationship between working memory and FL reading comprehension in our research represents the true relationship without being confounded. In summary, study one extended the first prediction in the Compensatory Encoding Model for L1 reading to FL reading. Whether FL readers are able to deploy compensatory use of reading strategies for word recognition inefficacy and working memory limitations will be investigated in study two.

## STUDY TWO

### Material and Methods

#### Participants

The participants in study two were 30 second year undergraduates (14 males and 16 females) recruited from the same university as in study one. They were between 18 and 23 years old with an average of 20.57 years ( $SD = 1.10$ ). Similarly, the average year of English learning for them was 7.5 years.

#### Materials

The materials used in study two were: the word recognition test, the working memory test, a reading passage for think-aloud, and the comprehension test. As the word recognition test and the

working memory test were exactly the same as in study one, please see study one for the details.

#### *The reading passage for think-aloud*

We used the text – *Ideas about Beauty* – to collect think-aloud data. This text was T3 used in the non-time constraint reading condition in study one. The reason for using a text from the non-time constraint reading condition was that the think-aloud method gave the readers ample time to process a text, simulating a non-time constraint reading condition.

Before think-aloud, the participants received detailed training in Chinese on how to verbalize their thoughts while reading and practiced think-aloud with a short expository text until they felt confident enough to carry out the think-aloud. To reduce demands on participants' verbal ability, they were free to articulate in either Chinese or English, or a combination of both (see Appendix 4 for the instructions). The think-aloud sessions were audio-recorded using a Lenovo audio-recorder.

#### *The reading comprehension test and scoring*

For the reading comprehension test, we used exactly the same 10 multiple choice questions (5 literal comprehension questions and 5 inferential comprehension questions) for *Ideas about Beauty* as in the study one. A correct answer for one question received one point (see study one for the detailed description and Appendix 3 for the text and the comprehension questions).

#### Data Collection Procedure

Data collection for study two took place in the students' free time in a quiet office. Each student first completed a think-aloud session followed by the reading test. They then completed the word recognition and working memory tests using a Dell laptop. The duration for data collection of study two ranged between 42 and 76 min.

#### Data Analysis

The data analysis in study two began with coding the think-aloud data. We first established a coding scheme on the basis of the two sources: Walczyk (2000) and Stevenson et al. (2007), in combination with the strategies from the data (see Appendix 5 for the coding scheme and examples). Reading strategies were categorized into: (1) language-oriented strategy, (2) content-oriented strategy, (3) re-reading above word-level, (4) pausing above word-level, and (5) meta-comment. Language-oriented strategies are directed toward "understanding the linguistic code of the text"; whereas content-oriented strategies are related to "building a mental model of the global conceptual content of the text" (Stevenson et al., 2007, p. 121). Re-reading above word-level refers to reprocessing part of the text which is more than a word. Pausing above word-level is defined as an interruption of decoding for 3 s or more of silence (Walczyk et al., 2004). Meta-comment is referred to as a reader's evaluation and reflection on his/her reading processes.

Language-oriented strategies comprised of translating, paraphrasing, grammatical problem-solving, discourse problem-solving, word processing problem-solving, and lexical inferencing. Translating occurs when part of the text is translated from English to Chinese. Paraphrasing is a strategy used to

rephrase meaning of one or parts of a sentence by replacing some words using synonyms, or reorganizing sentence structures either in English or Chinese. Grammatical problem-solving is used to disambiguate grammatical issues, while discourse problem-solving is defined as solving referential and/or cohesive devices problems. Word processing problem-solving involves possible strategies for solving inefficient word recognition, namely pausing and sounding out at word-level or below. Lexical inferencing strategies are used to make an informed guess of the meaning of an unknown word (Oxford and Scarcella, 1994; Kuhn and Stahl, 1998; Fraser, 1999; Paribakht and Wesche, 1999; Nassaji, 2003b; Bengelil and Paribakht, 2004).

Content-oriented strategies consisted of summarizing, interpreting, predicting, as well as questioning. Summarizing refers to recapping the gist of more than one sentence. Interpreting means integrating textual information with one's background knowledge to arrive at inferences of contents and purposes of part or all of a text. Predicting is defined as foreseeing the contents ahead, and questioning means raising questions concerning concepts conveyed in a text.

We used episodes as the unit of coding to code the frequency of reading strategy use with the assistance of the NVivo 9.2, which allowed audio files to be coded directly and the coded data were transformed into Excel for subsequent analysis. An episode was defined as a period when a reader "is unbrokenly occupied with the same component process" (Stevenson, 2005, p. 150). An episode ends at either the start of a different reading strategy or at a long pause of 10 s or more (Stevenson et al., 2007). The inter-coder reliability was calculated on randomly selected 6 think-aloud data, accounting for 20% of the total data. The Cohen's kappa was 0.79 for language-oriented strategies, 0.86 for content-oriented strategies, 0.88 for re-reading above word-level, 0.77 for pausing above word-level, and 0.79 for meta-comment. The frequency of reading strategy use was then correlated with word recognition, working memory, and reading comprehension to answer the research question of study two.

## Results

Altogether we identified a total of 2,339 strategies with an addition of seven unintelligible verbalizations, which were excluded from the analyses. The descriptive statistics of the five main types of strategy are displayed in Table 5, which shows that on average, our participants used 78 strategies to complete the reading task but the actual number of strategies used by the participants varied considerably as shown by a large *SD* of 30.10.

**TABLE 5 | Frequency of the five types of reading strategy.**

Types of reading strategy	Minimum	Maximum	<i>M</i>	<i>SD</i>
Language-oriented strategy	3	52	25.20	10.98
Content-oriented strategy	2	27	12.03	6.52
Re-reading above word-level	2	30	12.90	6.79
Pausing above word-level	0	15	3.37	3.94
Meta-comment	0	18	5.37	5.52

**TABLE 6 | Correlation analyses between reading strategy use, word recognition, working memory, and reading comprehension.**

Reading strategy use	Word recognition	Working memory	Comprehension
Language-oriented strategy	0.71**	−0.50**	0.77**
Content-oriented strategy	0.33	−0.33	0.20
Re-reading above word-level	0.40**	−0.57**	0.54**
Pausing above word-level	0.53**	−0.35	0.43**
Meta-comment	0.33	−0.24	0.33

\*\* $p < 0.01$  (two-tailed).

A one-way repeated ANOVA and Bonferroni *post hoc* analysis were employed to examine if there were significant differences of strategy use. The ANOVA showed that the participants differed significantly in terms of frequency of strategy use,  $F(4,29) = 66.21$ ,  $p < 0.01$ ,  $\eta_p^2 = 70$ . Bonferroni pair-wise comparison indicated that among the five main categories of reading strategy, the participants employed language-oriented strategies most frequently ( $M = 25.20$ ,  $SD = 10.98$ ). They used content-oriented strategies ( $M = 12.03$ ,  $SD = 10.98$ ) and re-reading above word-level ( $M = 12.90$ ,  $SD = 6.79$ ) less than half as frequently as language-oriented strategies. But there was no significant difference between content-oriented strategies and re-reading above word-level. The least frequently used reading strategies were pausing above word-level ( $M = 3.37$ ,  $SD = 3.94$ ) and meta-comment ( $M = 5.37$ ,  $SD = 5.52$ ), which did not differ from each other.

Table 6 presents the results of correlation analyses between word recognition, working memory, frequency of five types of reading strategy use, and reading comprehension. The relationship between word recognition and different kinds of reading strategies indicated that word recognition was significantly and positively correlated with language-oriented strategy use ( $r = 0.71$ ,  $p < 0.01$ ), re-reading above word-level ( $r = 0.40$ ,  $p = 0.03$ ), and pausing above word-level ( $r = 0.53$ ,  $p < 0.01$ ); whereas the association between word recognition and content-oriented strategy use ( $r = 0.33$ ,  $p = 0.08$ ) and meta-comment ( $r = 0.33$ ,  $p = 0.24$ ) were non-significant. This means that the students who had longer time to recognize English words (i.e., slower in word recognition) were associated with using more language-oriented strategy, re-reading above word-level, and pausing above word-level.

Resembling the results between word recognition and types of reading strategy use, we found that working memory did not significantly relate to content-oriented strategy use ( $r = -0.33$ ,  $p = 0.07$ ) and meta-comment ( $r = -0.24$ ,  $p = 0.21$ ), but it had significant and moderate association with language-oriented strategy use ( $r = -0.50$ ,  $p < 0.01$ ) and re-reading above word-level ( $r = -0.57$ ,  $p < 0.01$ ), suggesting that readers with a smaller working memory tended to use language-oriented strategy and paused more frequently. The correlation between working memory and pausing above word-level was also not significant ( $r = -0.35$ ,  $p = 0.06$ ).

Furthermore, we found that reading comprehension was significantly and positively correlated with language-oriented



strategy use ( $r = 0.77$ ,  $p < 0.01$ ), re-reading above word-level ( $r = 0.54$ ,  $p < 0.01$ ), and pausing above word-level ( $r = 0.43$ ,  $p = 0.02$ ). However, neither the association between reading comprehension and content-oriented strategy use ( $r = 0.20$ ,  $p = 0.30$ ) nor the association between reading comprehension and meta-comment ( $r = 0.33$ ,  $p = 0.08$ ) was significant.

## Discussion

The interrelationship between word recognition, working memory, strategy use, and reading comprehension demonstrated that when FL readers had sufficient time to process a reading task at hand, the more frequently they adopted language-oriented strategy, re-reading above word-level, and pausing above word-level, the better comprehension they achieved, and the reading comprehension was not affected by the performance of word recognition and working memory capacity. On the other hand, the content-oriented strategy use and meta-comment did not function as compensation for word recognition inefficiency and limited working memory. As described in the Compensatory Encoding Model, the kinds of strategies which are compensatory are those ones which mostly focus on solving language problems, re-reading, and pausing (Walczyk, 2000), just as what we found in our study, rather than strategies which focus on helping with understanding the global comprehension (content-oriented) and focus on the evaluation of their reading processes and behaviors (meta-comment).

Study two shows that the participants had a wide repertoire of reading strategies to draw on during FL reading, though the frequency of different types of reading strategies varied. Among five main categories of strategies classified, three of them, namely language-oriented strategy, re-reading above word-level, and pausing above word-level, were similar to the strategies with compensatory nature described in the Compensatory Encoding Model (Walczyk, 2000).

Regarding the relationship between word recognition, working memory, and use of reading strategies, our results of study two lent support to the second prediction of the Compensatory Encoding Model that when readers read without time restriction, word recognition efficiency and working memory tend to adversely relate to frequency of reading strategy use among Chinese college ELLs. This means that FL readers also displayed similar manner of using strategies to L1 readers to compensate for inefficiency in word recognition and limited working memory capacity so that their comprehension was not affected when they read in non-time constraint reading. However, we found that word recognition and working memory did not relate to all kinds of reading strategies, and content-oriented strategies and meta-comments seemed not to be compensatory. This finding was in fact in line with the Compensatory Encoding Model, in which the compensatory mechanisms are predominantly used to solve language problems rather than to direct toward understanding global conceptual comprehension (Walczyk and Taylor, 1996; Walczyk, 2000; Walczyk et al., 2001, 2007).

A possible reason why re-reading and pausing appeared to be effective compensatory strategies in FL reading comprehension could be explained by Kintsch's (1998) construction-integration perspective of comprehension. When readers' word recognition is not efficient enough, the initial derivation of propositional meanings from these processes tends to take up many working memory resources to construct a text-base of comprehension (the construction phase). By the end of the construction phase, readers' working memory may be exhausted, and this in turn leaves them with insufficient working memory resources to construct a situation-base of comprehension (the integration phase). In order for readers to have enough working memory in the integration phase, they may need to pause and re-read often in order to enable integration to take place. The reason that language-oriented strategies rather than content-oriented strategies were compensatory in FL text comprehension could be that the global conceptual understanding in reading largely depends on a reasonable understanding of the local meaning construction of the text (Kintsch, 1994, 1998). Thus, without achieving a good comprehension of phrases and sentences at the local level, it would be hard for FL learners to apply strategies toward building a coherent representation of the text.

The results of study two corroborated with studies in L1 reading with both children and adults (e.g., Walczyk and Taylor, 1996; Walczyk et al., 2001, 2004, 2007). The findings were also consistent with those in Stevenson's (2005) study. Stevenson also reported that among Dutch middle school ELLs, speed of recognizing English words was significantly related to language-oriented strategies but not to content-oriented strategies, and levels of reading comprehension. The similar results between the two studies suggest that language distance between L1 and FL in terms of orthography does not affect readers' successful use of reading strategies – a higher-order process in reading – to compensate for inefficient word recognition – a lower order process in reading. This appears to support Taylor and Taylor's (1995) assertion that processing strategies at the word level tend to be affected by differing orthographies across languages, but reading strategies at the higher-order comprehension processes tend to remain similar.

## Conclusion

To sum up, our research provided empirical evidence for the support of the Compensatory Encoding Model among Chinese ELLs. Despite that Chinese ELLs' English word recognition skills tend to be adversely affected by their holistic approach used in Chinese word recognition (Muljani et al., 1998; Akamatsu, 2003, 2005; Wang and Koda, 2005; Hamada and Koda, 2008, 2010), their compensatory reading strategy use (i.e., higher order skills) resembles that of native English speakers (e.g., Walczyk and Taylor, 1996; Walczyk et al., 2001, 2004) and ELLs whose L1 is also an alphabetic language (Stevenson, 2005). Our research indicates that rather than supporting the popular beliefs that word recognition efficiency is important in FL reading comprehension (Koda, 1996; Segalowitz, 2000), we found that

higher-order metacognitive strategy use is more important and predictive in FL reading comprehension, especially when readers are given ample time to process a reading task. As we found that language-oriented strategies are useful compensatory strategies, FL reading educators may wish to design training programs to teach students to use these strategies in order to facilitate text comprehension. The findings that contributions made by word recognition and working memory to FL reading comprehension are affected by reading varying time imply that time-restricted FL reading comprehension assessments may limit students' opportunities to apply some strategies to compensate for word processing inefficiency and cognitive resource limitation. Therefore, FL reading assessments may need to include multiple ways to measure FL readers' reading comprehension ability.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of 'National Statement on Ethical Conduct in Human Research, the Human Research Ethics Committee of the University of Sydney' 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 Human Research Ethics Committee of the University of Sydney.'

## REFERENCES

- Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Lang. Learn.* 53, 207–231. doi: 10.1111/1467-9922.00216
- Akamatsu, N. (2005). "Effects of second language reading proficiency and first language orthography on second language word recognition," in *Second Language Writing Systems*, eds V. Cook and B. Bassetti (Clevedon: Multilingual Matters), 238–259.
- Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Appl. Psycholinguist.* 29, 175–193. doi: 10.1017/S0142716408080089
- Alderson, J. (2000). *Assessing Reading*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511732935
- Alptekin, C. (2006). Cultural familiarity in inferential and literal comprehension in L2 reading. *System* 34, 494–508. doi: 10.1016/j.system.2006.05.003
- Alptekin, C., and Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *J. Res. Read.* 33, 206–219. doi: 10.1111/j.1467-9817.2009.01412.x
- Alptekin, C., and Erçetin, G. (2011). The effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading. *TESOL Q.* 45, 235–266. doi: 10.5054/tq.2011.247705
- Baddeley, A. D. (2006). "Working memory: an overview," in *Working Memory and Education*, ed. S. Pickering (Burlington, MA: Academic Press), 1–31. doi: 10.1016/b978-012554465-8/50003-x
- Baddeley, A. D. (2007). *Working Memory, Thought and Action*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780198528012.001.0001
- Baker, L., and Brown, A. L. (1984). "Metacognitive skills and reading," in *Handbook of Reading Research*, eds P. D. Pearson, R. Barr, M. L. Kamil, and P. Mosenthal (White Plains, NY: Longman), 353–394.
- Bengeleil, N., and Paribakht, T. S. (2004). L2 reading proficiency and lexical inferencing by university EFL learners. *Can. Mod. Lang. Rev.* 61, 225–249. doi: 10.3138/cmlr.61.2.225
- Brantmeier, C. (2005). Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension in Spanish. *Mod. Lang. J.* 89, 37–53. doi: 10.1111/j.0026-7902.2005.00264.x
- Carver, R. (1997). Reading for one second, one minute, or one year from the perspective of raiding theory. *Sci. Stud. Read.* 1, 3–43. doi: 10.1207/s1532799xssr0101\_2
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychon. Bull. Rev.* 12, 769–786. doi: 10.3758/BF03196772
- Forster, K. I., and Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124. doi: 10.3758/BF03195503
- Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Stud. Second Lang. Acquis.* 21, 225–241.
- Fukink, R. G., Hulstijn, J., and Simis, A. (2005). Does training in second-language word recognition skills affect reading comprehension? An experimental study. *Mod. Lang. J.* 89, 54–75. doi: 10.1111/j.0026-7902.2005.00265.x
- Grabe, W. (2009). *Reading in a Second Language: Moving from Theory to Practice*. New York, NY: Cambridge University Press.
- Grabe, W., and Stoller, F. L. (2011). *Teaching and Researching Reading*, 2nd Edn. Harlow: Pearson.
- Hamada, M., and Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Lang. Learn.* 58, 1–31. doi: 10.1111/j.1467-9922.2007.00433.x
- Hamada, M., and Koda, K. (2010). The role of phonological decoding in second language word-meaning inference. *Appl. Linguist.* 31, 513–531. doi: 10.1093/applin/amp061
- Harrington, M., and Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Stud. Second Lang. Acquis.* 14, 25–38. doi: 10.1017/S0272263100010457
- Haynes, M., and Carr, T. H. (1990). "Writing system background and second language reading: a component skills analysis of English reading by native speaker-readers of Chinese," in *Reading and its Development: Component Skills Approaches*, eds T. H. Carr and B. A. Levy (San Diego, CA: Academic Press), 375–421.

## AUTHOR CONTRIBUTIONS

The author confirms that she contributed to this paper by: Contributing substantially to the conception of the work; and the acquisition, analysis, and interpretation of the data; Drafting the work and revising it critically for important intellectual content; Approving the final version of the paper to be published; Agreeing to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## ACKNOWLEDGMENTS

The author wishes to acknowledge the financial support of the University of Sydney and New South Wales Institute for Educational Research. The author also wishes to express her sincere gratitude to Dr. Marie Stevenson for her kind guidance throughout the research.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00681/full#supplementary-material>

- Horiba, Y. (2000). Reader control in reading: effects of language competence, text type and task. *Discourse Process.* 29, 223–267. doi: 10.1207/S15326950dp2903\_3
- Hu, G. (2005). English language education in China: policies, progress, and problems. *Lang. Policy* 4, 5–24. doi: 10.1007/s10993-004-6561-7
- Jenkins, J. R., Johnson, E. S., and Hileman, J. (2004). When is reading also writing: sources of individual differences on the new reading performance assessments. *Sci. Stud. Read.* 8, 125–151. doi: 10.1207/s1532799xssr0802\_2
- Johnson, E. S., Jenkins, J. R., and Jewell, M. (2005). Analysing components of reading on performance assessment: an expanded simple view. *Read. Psychol.* 26, 267–283. doi: 10.1080/02702710590962587
- Juffs, A., and Harrington, M. W. (2011). Aspects of working memory in L2 learning. *Lang. Teach. Rev. Stud.* 42, 137–166. doi: 10.1017/S026144481000509
- Kendall, J. R., Mason, J. M., and Hunter, W. (2001). Which comprehension? Artifacts in the measurement of reading comprehension. *J. Educ. Res.* 73, 233–236. doi: 10.1080/00220671.1980.10885242
- Kintsch, W. (1994). “Discourse processing,” in *International Perspective in Psychological Science*, Vol. 2, eds G. d’Ydewalle, P. Eelen, and P. Bertelson (Hillsdale, NJ: Erlbaum), 135–155.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. New York, NY: Cambridge University Press.
- Koda, K. (1994). Second language reading research: problems and possibilities. *Appl. Psycholinguist.* 15, 1–28. doi: 10.1017/S0142716400006950
- Koda, K. (1996). L2 word recognition research: a critical review. *Mod. Lang. J.* 80, 450–460. doi: 10.1111/j.1540-4781.1996.tb05465.x
- Koda, K. (2000). Cross-linguistic variations in L2 morphological awareness. *Appl. Psycholinguist.* 21, 297–320. doi: 10.1017/S0142716400003015
- Koda, K. (2005). *Insights into Second Language Reading: A Cross-Linguistic Approach*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139524841
- Koda, K. (2007). Reading and language learning: cross-linguistic constraints on second language reading development. *Lang. Learn.* 57, 1–44. doi: 10.1111/0023-8333.101997010-11
- Kuhn, M., and Stahl, S. (1998). Teaching students to learn word meanings from context: a synthesis and some questions. *J. Lit. Res.* 30, 119–138. doi: 10.1080/10862969809547983
- LaBerge, D., and Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. doi: 10.1016/0010-0285(74)90015-2
- Leeser, M. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: topic familiarity and working memory. *Lang. Learn.* 57, 229–270. doi: 10.1111/j.1467-9922.2007.00408.x
- Malarcher, C. (2005). *College Reading Workshop*, 2nd Edn. Sachse, TX: Compass.
- Muljani, M., Koda, K., and Moates, D. (1998). Development of L2 word recognition: a connectionist approach. *Appl. Psycholinguist.* 19, 99–114. doi: 10.1017/S0142716400010602
- Nassaji, H. (2003a). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *Mod. Lang. J.* 87, 261–276. doi: 10.1111/1540-4781.00189
- Nassaji, H. (2003b). L2 vocabulary learning from context: strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Q.* 37, 645–670. doi: 10.2307/3588216
- Nassaji, H., and Geva, E. (1999). The contribution of phonological and orthographic processing skills to adult ESL reading: evidence from native speakers of Farsi. *Appl. Psycholinguist.* 20, 241–267. doi: 10.1017/S0142716499002040
- Nation, I. S. P. (2004). “A study of the most frequent word families in the British National Corpus,” in *Vocabulary in a Second Language*, eds P. Bogaards and B. Laufer (Amsterdam: John Benjamins), 3–13.
- Oxford, R., and Scarcella, R. (1994). Second language vocabulary learning among adults: state of the art in vocabulary instruction. *System* 22, 231–243. doi: 10.1016/0346-251X(94)90059-0
- Paribakht, T. S., and Wesche, M. (1999). “Incidental” vocabulary acquisition through reading: an introspective study. *Stud. Second Lang. Acquis.* 21, 203–220. doi: 10.1080/02643290500538315
- Perfetti, C. A. (1985). *Reading Ability*. New York, NY: Oxford University Press.
- Perfetti, C. A. (1988). “Verbal efficiency theory in reading ability,” in *Reading Research: Advances in Theory and Practice*, eds M. Daneman, G. E. MacKinnon, and T. G. Waller (New York, NY: Academic Press), 109–143.
- Perfetti, C. A. (2007). Reading ability: lexical quality to comprehension. *Sci. Stud. Read.* 11, 357–383. doi: 10.1080/10888430701530730
- Phakiti, A. (2007). *Strategic Competence and EFL Reading Test Performance: A Structural Equation Modeling Approach*. Frankfurt: Peter Lang.
- Pressley, M., and Afflerbach, P. (1995). *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading*. Hillsdale, NJ: Erlbaum.
- Rai, M., Loschky, L. C., Harris, R. J., Peck, N. R., and Cook, L. G. (2011). Effects of stress and working memory capacity on foreign language readers’ inferential processing during comprehension. *Lang. Learn.* 61, 187–218. doi: 10.1111/j.1467-9922.2010.00592.x
- Segalowitz, N. (2000). “Automaticity and attentional skill in fluent performance,” in *Perspectives on Fluency*, ed. H. Riggenbach (Ann Arbor, MI: University of Michigan Press), 200–219.
- Segalowitz, N. (2003). “Automaticity and second language acquisition,” in *The Handbook of Second Language Acquisition*, eds C. Doughty and M. Long (Oxford: Blackwell), 382–408. doi: 10.1002/9780470756492.ch13
- Segalowitz, N., and Hulstijn, J. (2005). “Automaticity in bilingualism and second language learning,” in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (Oxford: Oxford University Press), 371–388.
- Seigneuric, A., Ehrlich, M., Oakhill, J., and Yuill, N. M. (2000). Working memory resources and children’s reading comprehension. *Read. Writ.* 13, 81–103. doi: 10.1023/A:1008088230941
- Service, E., Simola, M., Metsanheimo, O., and Mauri, S. (2002). Bilingual working memory span is affected by language skill. *Eur. J. Cogn. Psychol.* 14, 383–408. doi: 10.1080/09541440143000140
- Shaw, P., and McMillion, A. (2008). Proficiency effects and compensation in advanced second-language reading. *Nord. J. Engl. Stud.* 7, 123–143.
- Shiotsu, T. (2009). “Reading ability and components of word recognition speed: the case of L1-Japanese EFL learners,” in *Second Language Reading Research and Instruction*, eds Z. Han and N. Anderson (Ann Arbor, MI: University of Michigan Press), 15–37.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychol. Bull.* 87, 245–251. doi: 10.1016/j.neuroimage.2010.08.042
- Stevenson, M. (2005). *Reading and Writing in a Foreign Language: A Comparison of Conceptual and Linguistic Processes in Dutch and English*. Doctoral dissertation, Universiteit van Amsterdam, Amsterdam.
- Stevenson, M., Schoonen, R., and de Glopper, K. (2007). Inhibition or compensation? A multi-dimensional comparison of reading processes in Dutch and English. *Lang. Learn.* 57(Suppl. 1), 115–154. doi: 10.1111/j.1467-9922.2007.00414.x
- Taylor, I., and Taylor, M. M. (1995). *Writing and Literacy in Chinese, Korean and Japanese*. Amsterdam: John Benjamins. doi: 10.1075/swll.3
- Tsai, Y. R. (2008). Lower-level processing and strategy use in L1 and L2 reading comprehension among Chinese-speaking EFL learners. *Lacus Forum* 34, 269–280.
- Unsworth, N., Heitz, R. P., Schrock, J. C., and Engle, R. W. (2005). An automated version of the operation span task. *Behav. Res. Methods* 37, 498–505. doi: 10.3758/BF03192720
- van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., et al. (2004). Linguistic knowledge, processing speed and metacognitive knowledge in first and second language reading comprehension: a componential analysis. *J. Educ. Psychol.* 96, 19–30. doi: 10.1037/0022-0663.96.1.19
- van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Snellings, P., Simis, A., et al. (2003). Linguistic knowledge, metacognitive knowledge and processing speed in L1, L2 and L3 reading: a structural equation modelling approach. *Int. J. Biling.* 7, 7–25. doi: 10.1177/13670069030070010201
- van Gelderen, A., Schoonen, R., Stoel, R., de Glopper, K., and Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: a longitudinal analysis of constituent components. *J. Educ. Psychol.* 99, 477–491. doi: 10.1037/0022-0663.99.3.477
- Walczyk, J. J. (1993). Are general resource notions still viable in reading research? *J. Educ. Psychol.* 85, 127–135. doi: 10.1037/0022-0663.85.1.127

- Walczyk, J. J. (1995). A test of the compensatory-encoding model. *Read. Res. Q.* 30, 396–408. doi: 10.2307/747623
- Walczyk, J. J. (2000). The interplay between automatic and control process in reading. *Read. Res. Q.* 35, 554–556. doi: 10.1598/RRQ.35.4.7
- Walczyk, J. J., Marsiglia, C., Bryan, K., and Naquin, P. (2001). Overcoming inefficient reading skills. *J. Educ. Psychol.* 93, 750–757. doi: 10.1037/0022-0663.93.4.750
- Walczyk, J. J., Marsiglia, C., and Johns, A. K. (2004). Children's compensations for poorly automated reading skills. *Discourse Process.* 37, 47–66. doi: 10.1207/s15326950dp3701\_3
- Walczyk, J. J., and Taylor, R. J. (1996). How do the efficiencies of reading subcomponents relate to looking back in text? *J. Educ. Psychol.* 88, 737–745. doi: 10.1037/0022-0663.88.3.537
- Walczyk, J. J., Wei, M., Griffith-Ross, D. A., Goubert, S. E., Cooper, A., and Zha, P. (2007). Development of the interplay between automatic processes and cognitive resources in reading. *J. Educ. Psychol.* 99, 867–877. doi: 10.1037/0022-0663.99.4.867
- Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Appl. Linguist.* 25, 315–339. doi: 10.1093/applin/25.3.315
- Wang, M., and Koda, K. (2005). Commonalities and differences in word identification skills among learners of English as a second language. *Lang. Learn.* 55, 71–98. doi: 10.1111/j.0023-8333.2005.00290.x
- Waters, G. S., and Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Q. J. Exp. Psychol.* 49, 51–79. doi: 10.1080/713755607
- Yamashita, J. (2013). Word recognition subcomponents and passage level reading in a foreign language. *Read. Foreign Lang.* 25, 52–71.

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Han. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Contribution of Word Reading Speed to Reading Comprehension in Brazilian Children: Does Speed Matter to the Comprehension Model?

Alessandra G. Seabra<sup>1\*</sup>, Natália M. Dias<sup>2</sup>, Tatiana Mecca<sup>2</sup> and Elizeu C. Macedo<sup>1</sup>

<sup>1</sup> Developmental Disorders Program, Universidade Presbiteriana Mackenzie, São Paulo, Brazil, <sup>2</sup> Educational Psychology Post-Graduation Program, Centro Universitário FIEO, Osasco, Brazil

## OPEN ACCESS

### Edited by:

Giseli Donadon Germano,  
Universidade Estadual Paulista –  
UNESP, Brazil

### Reviewed by:

Maria Regina Maluf,  
Pontifícia Universidade Católica de  
São Paulo (PUCSP), Brazil  
Grazielle Kerges Alcantara,  
University of São Paulo, Brazil

### \*Correspondence:

Alessandra G. Seabra  
alessandragseabra@gmail.com

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 25 August 2015

**Accepted:** 05 April 2017

**Published:** 20 April 2017

### Citation:

Seabra AG, Dias NM, Mecca T  
and Macedo EC (2017) Contribution  
of Word Reading Speed to Reading  
Comprehension in Brazilian Children:  
Does Speed Matter to the  
Comprehension Model?  
Front. Psychol. 8:630.  
doi: 10.3389/fpsyg.2017.00630

Studies have suggested that reading speed (RS) or fluency should be a component of reading comprehension (RC) models. There is also evidence of a relationship between RS and RC. However, some questions remain to be explored, as the changes in such a relationship may be a function of development. In addition, while there are studies published with English speakers and learners, less evidence exists in more transparent orthographies, such as Portuguese. This study investigated the relationship between RC and RS in typical readers. Objectives included elucidating the following: (1) the contribution of RS to RC controlling for intelligence, word recognition, and listening and (2) the differential relationships and contributions of RS to comprehension in different school grades. The sample of participants comprised 212 students ( $M = 8.76$ ;  $SD = 1.06$ ) from 2nd to 4th grade. We assessed intelligence, word recognition, word RS, listening, and RC. Performance in all tests increased as a function of grade. There were significant connections between RC and all other measures. Nonetheless, the regression analysis revealed that word RS has a unique contribution to RC after controlling for intelligence, word recognition, and listening, with a very modest but significant improvement in the explanatory power of the model. We found a significant relationship between RS and RC only for 4th grade and such relationship becomes marginal after controlling for word recognition. The findings suggest that RS could contribute to RC in Portuguese beyond the variance shared with listening and, mainly, word recognition, but such a contribution was very small. The data also reveal a differential relationship between RS and RC in different school grades; specifically, only for the 4th grade does RS begins to relate to RC. The findings add a developmental perspective to the study of reading models.

**Keywords:** reading competence, fluency, learning, cognitive assessment, cognitive models

## INTRODUCTION

The National Reading Panel highlights three important areas for reading, learning and competence: alphabets, which is related to word recognition skills; fluency, the ability to read with speed, accuracy, and proper expression (prosody); and comprehension, here understood as reading comprehension (RC) or reading competence, a complex process that integrates other abilities such

vocabulary comprehension and strategies for understanding. In this statement by the National Reading Panel, it is clear that the notion of independent yet correlated components exists and recognizes that fluency is related to word recognition skills, but its position is clear about the independence of such processes — that is, word recognition does not necessarily lead to fluency (National Reading Panel (US) and National Institute of Child Health and Human Development (US), 2000). A relatively recent review on the theme-compiled studies in the area has corroborated the view of fluency as a concept that integrates accuracy, automaticity, and prosody in oral reading (Kuhn et al., 2010).

This componential view of reading is also represented in cognitive models such as the Simple View of Reading (SVR), a model proposed by Gough and Tunmer (1986) that suggests that reading competence (R) can be explained by two components, decoding (D) and listening comprehension (LC), which is expressed by the formula  $R = D \times LC$ . A modified and elaborated version is the component model of reading (CMR; Aaron et al., 2008). CMR postulates that three domains can impact reading learning and competence: (1) the cognitive domain, comprising word recognition and comprehension (similar to SVR, but note that the “word recognition” component expands the idea of decoding); (2) the psychological domain, which includes motivation, interest, learning styles and other psychological phenomena; and (3) the ecological domain, including variables such as culture, home and classroom environments.

Considering the cognitive components (SVR or cognitive domain of CMR), evidence already suggests that decoding and LC can explain RC in different orthographies. For example, with students from 2nd, 3rd, and 4th grades, approximately 60% of the variance in RC was explained by decoding and listening for Spanish speakers, while approximately 50% of the variance in RC was explained for English speakers. Spanish is considered as having a transparent orthography due to the close relationship between phonemes and graphemes. On the other hand, English has an opaque orthography, in which there is a lack of one-to-one correspondence between sound and the letter. Another interesting finding shows that the contribution of components to RC in Spanish 3rd graders and English 4th graders was very similar. Authors interpreted this result as Spanish-speaking children mastering decoding skills earlier than English-speaking children, who use a more opaque orthography. The same study investigated the model in Chinese orthography, which has morphosyllabic orthography (in which a character represents a word). Some similar patterns were found, with component skills accounting for 31% of the variance in Chinese RC at 2nd grade and 42% at 4th grade. Despite this, the explanatory power of the model for Chinese was smaller than in English and in Spanish (Joshi et al., 2012).

Expanding this view and considering the suggestion of an additional component to the SVR model by Joshi and Aaron (2000), speed has also been studied as a new and independent skill that contributes to reading competence. Indeed, in the original work of Joshi and Aaron (2000) with 3rd graders, D and LC accounted for 48% of variance in RC, and the addition of the speed component, which was assessed in a letter-naming task, added 10% to the explanatory power of the model, suggesting

some unique contribution of the speed component. More recently, Aaron et al. (2008) supported word recognition and linguistic comprehension as the main components of cognitive domains of reading (explaining 37 to 41% of RC from 2nd to 5th grade), while speed (that authors referred as processing speed, assessed by a letter-naming task) shows some inconsistent contributions that can vary from 11% in 2nd grade to 2.5% in 5th grade. Such findings suggest that there is a decreasing trend in the effect of speed on RC with development. Another study with 4th and 5th graders also found some significant and unique contribution from speed to RC, as assessed in a picture-naming task, but such a contribution was small, varying from 2% for 4th grade and 1.4% for 5th grade. Authors argue that such a low contribution is due to the speed that has already had an effect upon decoding (Johnston and Kirby, 2006).

Other studies had investigated the same question — that is, the cognitive components or contributors of competent reading — but used a different measurement than speed, which was reading fluency itself. According to Kuhn et al. (2010), reading fluency combines accuracy, automaticity, and prosody. In fact, studies have investigated one or more of these aspects of fluency in different units of reading, such as words or texts, with some contradictory results.

With a sample of 5th grade students from 11 to 12 years of age, Turkyilmaz et al. (2014) found that, despite all fluency measures that were significantly related with RC, oral reading fluency had the strongest contribution to prediction of RC when compared to silent reading fluency and retell fluency. Similarly, Klauda and Guthrie (2008) investigated the relationships between three measures of fluency (at the word, syntactic – phrase and sentence units of text, and passage levels) and RC in 5th grade students. The authors found that three types of fluency were individually related to performance on a RC test. In addition, some evidence suggested that RC and reading fluency at the syntactic level appeared to have a bidirectional relationship. The authors suggest that fluency and comprehension becomes more similar over time up to the age of approximately 10 or 12 years. In this sense, they argue that it may be useful to examine the existence of different relationships between fluency (and different fluency skills) and comprehension across different grade levels. Furthermore, a national (Brazilian) study that evaluates prosody found no significant relationship with student RC from 3rd to 5th grades and only marginal non-significant trends in this relationship for 3rd and 4th grade students (Martins and Capellini, 2014).

Despite the relevance of fluency for reading competence (National Reading Panel (US) and National Institute of Child Health and Human Development (US), 2000; Hudson et al., 2005), studies in the area show that there is less agreement as to how to assess such skills, as some studies assessed naming speed (Joshi and Aaron, 2000; Johnston and Kirby, 2006; Aaron et al., 2008), some assessed word or different units of text reading rates (Klauda and Guthrie, 2008; Turkyilmaz et al., 2014), and others assess prosody (Martins and Capellini, 2014). This fact can lead to difficulty in making comparisons between the findings in the area because there is no clarity regarding the independence or overlap between these several measurements. Additionally, aspects of

reading fluency had been poorly studied in Brazil (Puliezi and Maluf, 2014). In this context, we evaluated accuracy and speed at the word level [that is, the reading of isolated words and word reading speed (RS), respectively], listening and RC at the sentence level. We choose these particular levels of analyses because the use of more complex levels could imply more demands on the tasks, such as working memory on comprehension of longer text, and also because we intend to investigate the relationship between such skills in children in the first years of elementary school, in which competence at the text level could not be consolidated.

Previous findings in Brazil elucidated the development of listening and RC of sentences, as well as word reading strategies. Dias et al. (2015) found that both listening and RC developed in 1st to 3rd grade, with no difference between 3rd and 4th grades in the tests. Regarding word reading strategies, the study suggested further development of alphabetical (decoding or phonological route) and logographic (contextual word recognition) strategies in early literacy (mainly in 1st and 2nd grades), and further development of orthographic processing in more advanced grades (3rd and mainly 4th grades). In this sense, one wonders if, with the proficiency in word recognition and development of orthographic reading, RS grows. Therefore, it can be expected that faster reading with educational progression is associated with the automation of alphabetic reading and use of orthographic reading and, consequently, with better RC. In this case, while in the early years of elementary school, word recognition would be the most important skill for RC, and with educational progression, fluency would become more relevant.

Therefore, more studies are needed and some questions remained to be further explored. For example, while there are studies that examined English middle-school student speakers/learners, less evidence exists in other orthographies, such as Portuguese, and studies of earlier grades. In this context, this study investigated the relationship between sentence RC and word RS in typical readers during elementary school. Specifically, we wished to elucidate (1) the contribution of RS to RC after controlling for intelligence, word recognition, and listening and (2) the differential relationships and contributions of RS to comprehension in different school grade levels. Our hypotheses are (1) word RS will present a modest but significant contribution to the explanatory model of RC, showing some unique contributions not accounted for by intelligence, word recognition, and listening and (2) more consistent relationships and contributions will be established between RS and comprehension in the earlier grades in our sample.

## MATERIALS AND METHODS

### Participants

The participants comprised 223 students from 2nd to 4th grade in São Paulo, Brazil. From this initial sample, 11 participants were excluded (10 with histories of academic failure and 1 with an indicator of intellectual disability as assessed by the Raven

test). The final sample comprised 212 students with 51.4% female (Mean age = 8.76 years;  $SD = 1.06$ ), and included 85 students from the 2nd grade (Mean age = 7.92;  $SD = 0.727$ ); 52 students from the 3rd grade (Mean age = 8.65;  $SD = 0.623$ ), and 75 students from the 4th grade (Mean age = 9.80;  $SD = 0.658$ ). In the final sample, there were no students with motor or sensorial disabilities that would impair their performance in the tests.

## Instruments

### Words and Non-words Reading Competence Test – WNw-RCT

The WNw-RCT (Seabra and Capovilla, 2010) assesses competence in reading isolated words and is comprised of 70 test items, each of which features a picture paired with a written word. There are seven different types of items: correct regular words [CR, e.g., the word 'FADA' (fairy in English) with the image of a fairy], correct irregular words [CI, e.g., the word 'BRUXA' (witch) with the image of a witch]; semantic changes [SC, e.g., the word 'RÁDIO' (radio in English) with the image of a phone]; visual changes [VC, e.g., the word 'TEIEUISÃO' (the correct spelling is TELEVISÃO) with the image of a television]; phonological changes [PC, e.g., the word 'MÁCHICO' (the correct spelling is MÁGICO) with the image of a wizard]; weird non-words [WN, e.g., the word 'MELOCE' (a word that does not exist in Portuguese)], and homophone non-words [HN, e.g., the word 'TACSI' (the correct spelling is TAXI) with the image of a taxi]. The children need to choose the corrected word and reject any semantic errors or non-words. Despite allowing for the differential assessment of reading strategies (logographic, alphabetic, and orthographic), we used the total score in this study, an index of word-recognition skills.

### Contrastive Test of Listening and Reading Comprehension – CTLRC

The CTLRC (Capovilla and Seabra, 2013) assesses listening and RC skills. The instrument consists of two subtests: RC and LC, each with 40 test items arranged in order of increasing difficulty. For each item, the child must choose between five alternative figures, the one that corresponds to the sentence heard in the case of the LC subtest or read in the case of the RC.

### Reading Speed Test – RST

Reading Speed Test (Montiel, 2008) was used to assess RS. The test requires the subject to read isolated words presented in the middle of the computer screen as quickly as possible. Test scores show the successes and mistakes and the time required for reading (speed measure). The RST consists of 60 items divided into four parts (P1 to P4): 15 irregular words (P1), 15 pseudo-words (P2), 15 words related to content (i.e., nouns \_ P3) and 15 words related to function (such as conjunctions, adjectives, and adverbs \_ P4). In this study, we used only parts 1 and 2 (thus, 30 items). All of the words containing three to four letters were presented using the Times New Roman font, size 72 in black ink, and for an indefinite time on the screen. Only the times for items resulting in 100% reading accuracy were considered in the analysis.

## Raven's Colored Standard Progressive Matrices – RCSPM

The RCSPM assess general intellectual ability (Angelini et al., 1999), specifically, the reasoning related to the formation of new creative insights and high-level functions.

## Procedure

Our study was approved by the Research Ethics Committee. Agreement Terms were sent to the students' parents, asking for their consent to carry out the research. The WNW-RCT and CTLRC were collectively applied in the classroom in three sessions, one for the WNW-RCT and one for each CTLRC subtest (Reading and LC), allowing for a 1-week interval between tests. Raven's Test and the RST were individually applied in two sessions of 30 min, one for each instrument. The assessment sessions lasted approximately 30 min each. The assessment occurred in the middle of the school year.

## Statistical Analysis

We performed descriptive and inferential (the ANOVA of the grade effect) statistics for each reading measurement. For groups with significant differences between the performances of the instruments, effect size (ES) analyses were conducted. To investigate the correlations between RC and other measures, we performed a Pearson correlation analysis (with the total sample). A hierarchical linear regression analysis was performed to investigate whether RS has some unique contribution (Model 3) to RC after controlling for intelligence (Model 1), word recognition, and listening (Model 2). To investigate differential correlations between RS and RC as a function of grade level, we performed a Pearson correlation and a partial correlation (controlling for word recognition) analysis independently for each grade level.

## RESULTS

**Table 1** presents descriptive statistics for each measure as a function of grade and for the total sample. Significant effects were found for all measures with increases in scores and speed (and a decrease in response time) as a function of grade. ES analyses found some important effects. For reading speed (RST), ES was moderate between the 2nd, 3rd ( $d = 0.62$ ) and 4th grade ( $d = 0.62$ ). For recognition of words and pseudo words (WNW-RCT), large ESs were found between 2nd and 3rd grade ( $d = 1.15$ ) and between 2nd and 4th grade ( $p = 1.44$ ). For reading comprehension (RC-CTLRC), large ESs were also verified between 2nd and 3rd grade ( $d = 0.97$ ) and 2nd and 4th grade ( $p = 1.12$ ). For listening (CTLRC-LC), large ESs were observed between 2nd and 3rd grade ( $d = 0.86$ ) and between 2nd and 4th ( $d = 0.82$ ).

**Table 2** presents the correlations between the measurements. RC had a positive significant relationship of a high magnitude with listening; of moderate magnitude with word recognition and a negative significant relationship of low magnitude with RS.

Based on the relationships found, we performed a hierarchical regression analysis with RC as a criterion variable. Three models

were generated, all with satisfactory fit ( $p < 0.001$ ). **Table 3** shows the models and coefficients for each predictive measure.

Model 1 includes only the intelligence measurement and explains 16% of the variance in RC. Model 2 includes the measurements of listening and word recognition, and the predictive power of the model increased to 52.6%. It is worth noting that with the inclusion of such variables, the contribution of intelligence is no longer significant. Model 3 adds the speed measurement. Although modest, the inclusion of this variable increased the explanatory power of the model in a significant way, and the contribution of RS for RC was significant, despite the control of previous variables, which suggests some unique contribution.

In addition, we explore the relationships between speed reading measurements and RC throughout the grades. As **Table 4** shows, there was a significant, negative, and low relationship between RC and speed only for the 4th grade, showing that children who needed more time in word reading (slower readers) tended to have lower RC. Based on such relationships, we performed a partial correlation analysis between RS and RC, controlling for word recognition. For the 4th grade, the relationship previously found became marginal. **Table 5** shows these results.

## DISCUSSION

The first objective of this study was to investigate the contribution of RS for RC after controlling for intelligence, word recognition, and listening. First, we conducted an ANOVA to verify the development of reading ability during the initial elementary school grades. Results revealed that listening and RC, word recognition, and RS developed during the 2nd, 3rd, and 4th grades, and no significant differences were found between the more advanced grades. These results suggest that, from the 2nd to 3rd grades, there is an important development of reading skills in Brazilian students, and there may be a consolidation in the progression from 3rd to 4th grade. Such differences in performance along grade levels were expected (e.g., Dias et al., 2015) and encouraged conducting some analyses separately for each grade level, as will be discussed later.

The correlation analysis between RC and listening, word recognition, and RS revealed significant relationships in all cases. Considering only reading abilities, as expected, listening and word recognition presented the strongest relationships with RC, while RS was only connected to low magnitude with comprehension. Expanding the results of this analysis, the findings from the regression analysis revealed that even with intelligence controlled, listening and word recognition can explain 52.6% of the variance in RC. This evidence corroborates the value of the components of the SVR (or the cognitive domain of CMR). Indeed, our data virtually replicate the previous findings of Joshi and Aaron (2000), who found that listening and word recognition accounted for 48% of variance in RC in 3rd graders, and Aaron et al. (2008), who found that these abilities explain approximately 40% of RC in students from 2nd to 5th grade.



**TABLE 1 | Descriptive and inferential statistics of grade effect on listening and reading measurements.**

		<i>N</i>	Mean	<i>SD</i>	<i>F</i> (2,211)	<i>p</i>	<i>Post-hoc analyses</i>
Word recognition	2	85	56.82	5.504	50.780	<0.001	2 < 3, 4
	3	52	62.40	3.560			
	4	75	64.49	5.134			
	Total	212	60.91	6.017			
Reading comprehension	2	85	27.13	11.025	37.153	<0.001	2 < 3, 4
	3	52	35.83	3.730			
	4	75	36.60	3.915			
	Total	212	32.61	8.800			
Listening comprehension	2	85	35.05	4.290	21.826	<0.001	2 < 3, 4
	3	52	38.10	1.660			
	4	75	38.00	2.594			
	Total	212	36.84	3.538			
Reading speed	2	85	0.7824	0.28911	12.579	<0.001	2 > 3, 4
	3	52	0.6324	0.13680			
	4	75	0.6448	0.09567			
	Total	212	0.6969	0.21434			

Word recognition: score in *WNw-RCT* – Words and Non-words Reading Competence Test; Reading comprehension: score in *CTLRC-RC* – Reading Comprehension subtest of the Contrastive Test of Listening and Reading Comprehension; Listening comprehension: score in *CTLRC-LC* – Listening Comprehension subtest of the Contrastive Test of Listening and Reading Comprehension; Reading speed: one word reading (locution) time (in seconds) in *RST* – Reading Speed Test.

**TABLE 2 | Correlation matrix between speed, word recognition, listening and reading comprehension (total sample).**

		Reading comprehension
Reading speed	<i>r</i>	−0.29
	<i>p</i>	0.000
Word recognition	<i>r</i>	0.58
	<i>p</i>	0.000
Listening comprehension	<i>r</i>	0.66
	<i>p</i>	0.000

In addition, Joshi et al. (2012), examining students from 2nd, 3rd, and 4th grades, found that listening and word recognition could explain 60% of the variance in RC for Spanish speakers, while approximately 50% of the variance in RC was explained for English speakers. In our sample, with Portuguese (Brazilian Portuguese) speakers, we found that 52.6% of variance in comprehension can be explained by the component skills. While Spanish has a transparent orthography, English is an opaque orthography. Portuguese has irregularities and rules but in general has a more transparent orthography than English. Despite this, our findings were more similar to results found in English rather than Spanish speakers.

The inclusion of RS in the regression increased the explanatory power of our model to 54.4% (an increase of 1.8%). Although low, the unique contribution of RS to comprehension was significant. Results in this area are debatable. For example, one study found that speed (assessed in a naming speed task) added 10% to the explanatory power of the RC model (Joshi and Aaron, 2000), while other evidence revealed contributions that can vary from 11% in 2nd grade to 2.5% in 5th grade (Aaron et al., 2008) and from 2% for 4th grade and 1.4% for 5th grade (Johnston and Kirby, 2006). Despite having different measures of speed, as used with word RS, our results are very

similar to Johnston and Kirby's (2006), as we also found a low contribution of naming speed (1.8%) for RC models. Studies involving reading fluency (instead of naming speed) also support the relationship between RC and measures of fluency (Klauda and Guthrie, 2008). For instance, Turkyilmaz et al. (2014) found that in 5th grade, fluency (including oral reading fluency, silent reading fluency, and retell fluency) explained 57% of the total variance in RC, but other predictors were not used in this study.

In this sense, our first hypothesis (word RS will present a modest but significant contribution to the explanatory model of RC, showing some unique contribution not accounted for by intelligence, word recognition or listening) proved correct; despite this, we expected the greatest contribution of word RS in our sample with students in the first years of elementary school. It is possible that we were not able to find greater contributions of speed for the model due the fact that speed may already have an effect upon word recognition; or, alternatively, the impact of RS on comprehension should not be relevant, for example, for these younger students, in which the decoding skill is so incipient that there is no variation among students. According to Yovanoff et al. (2005), fluency is a more meaningful measurement when the variation among student reading fluency is maximized. This hypothesis can be elucidated by our second objective, which is to investigate the differential relationships and contributions of RS to comprehension in different school grade levels.

We found a significant relationship between RS and RC only for the 4th grade. Furthermore, when controlling for word recognition, such a relationship was only a marginal trend. No significant relationships were found for 2nd and 3rd grades. The data reveal that RS only begins to be related to reading competence in 4th grade, with no correlation in early stages of elementary school. In this sense, the results corroborated our second hypothesis, which is that faster reading is expected with academic progression.

**TABLE 3 | Models from the regression analysis of the prediction of reading comprehension.**

Model		Beta	t	p	R <sup>2</sup>	R <sup>2</sup> adjusted
1	(Constant)		16.712	0.000	0.167	0.163
	Intellectual ability	0.408	6.482	0.000		
2	(Constant)		−7.910	0.000	0.533	0.526
	Intellectual ability	0.090	1.661	0.098		
	Word recognition	0.339	6.300	0.000		
	Listening comprehension	0.458	7.887	0.000		
3	(Constant)		−5.776	0.000	0.552	0.544
	Intellectual ability	0.100	1.882	0.061		
	Word recognition	0.305	5.643	0.000		
	Listening comprehension	0.446	7.823	0.000		
	Reading speed	−0.144	−2.989	0.003		

Intellectual ability: percentile in Raven's Colored Standard Progressive Matrices.

**TABLE 4 | Correlation matrix between reading comprehension and reading speed in each grade level.**

School grade	2	r	−0.17
		p	0.122
	3	r	0.12
		p	0.409
	4	r	<b>−0.30</b>
		p	<b>0.009</b>

Significant relations highlighted in bold.

**TABLE 5 | Partial correlation matrix between reading comprehension and reading speed in each grade level, controlling for word recognition ability.**

School grade	2	r	−0.14
		p	0.198
	3	r	0.10
		p	0.492
	4	r	<b>−0.21</b>
		p	<b>0.069</b>

Controlling for Word recognition (score in WNW-RCT). Significant relations highlighted in bold.

Regarding this finding, it is interesting to note that Aaron et al. (2008) suggest a trend for the decrease in the effect of speed with academic progression. Additionally, Klauda and Guthrie (2008) hypothesize that word-level fluency could be more connected with comprehension at the beginning of elementary school. Our results indicated the opposite. Klauda and Guthrie (2008) indeed claimed that fluency and comprehension should become more similar (then, more correlated) over time up to the age of 10 or 12, which is the mean age of our 4th grade students. Furthermore, we can use developmental data to explain our results. For example, previous research with Brazilian children suggests that word recognition is better developed only at 4th grade, with more use of orthographic strategy; that is, automaticity (and then speed) may be more important for these students. On the other hand, students from 3rd grade, but mainly 2nd grade, are very dependent on decoding skills, which is a slower process, and at this initial stage of learning, accuracy could be more important

than speed to comprehension (Dias et al., 2015). Therefore, the significant relationship in this study between RS and RC only for the 4th grade can be explained because, in these more advanced grades, the importance of alphabetic reading automation and the use of orthographic reading automation increase.

Additionally, according to our results, at least 45% of the variance in RC remains unexplained. In this sense, more studies are needed to clarify other demands of the reading competence model. Our study had some limitations, including the small number of participants, which makes the performance of regression analyses separated by grade level inviable to. Additionally, another limitation concerns the study design, of cohorts of different grade levels, instead a longitudinal study. This design prevents sure whether the observed differences in the relevance of RS for RC are, in fact, due to the child's developmental stages. Longitudinal studies should be conducted to ensure that the differences observed here are not due to specific characteristics of the study sample. Another limitation concerns the fluency test. We used a word speed test. However, fluency has been defined as a combination between accuracy, automaticity, and prosody (Kuhn et al., 2010); therefore, other tests, such as text RS, can bring different results.

On the other hand, one strength of the current study concerns the language/orthography feature. As most studies on SVR or reading competence models and components are performed with English speaking participants, we expand these data for a different orthography in working with Portuguese speaking students. In conclusion, some educational implications may be extrapolated as the comprehension of cognitive models and their components provide some framework when studying and diagnosing reading difficulties, providing professionals with guidelines about what skills to evaluate and how to plan interventions for children with reading difficulties.

## FINAL VIEWS

Findings suggest that RS can contribute to RC beyond the variance shared with listening and word recognition, but such contributions were low. The data also reveal a differential

relationship and contribution of RS in different school grades. Specifically, only in the 4th grade does RS begin to have some association to reading competence. The findings add a developmental perspective to the study of reading models and expand the previous research on reading components and models to a more transparent orthography, such as Portuguese.

## ETHICS STATEMENT

The project was approved by the Ethics Committee of the Universidade São Francisco, Brazil. Upon approval, the person legally responsible for the child signed the free and informed consent term, according to the rules of the National Council of the Brazil Health. In addition, children expressed

verbally their consent to participate. Participation in the study did not offer risks and participants could withdraw at any time.

## AUTHOR CONTRIBUTIONS

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

## FUNDING

This work was supported by grant for the AGS from CNPq (No. 309625/2013-0) and ECM from CNPq (No. 309453/2011-9).

## REFERENCES

- Aaron, P. G., Joshi, R. M., Gooden, R., and Bentum, K. E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading: an alternative to the discrepancy model of LD. *J. Learn. Disabil.* 41, 67–84. doi: 10.1177/0022219407310838
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., and Duarte, J. L. (1999). *Matrizes Progressivas Coloridas*. São Paulo: CETEPP.
- Capovilla, F. C., and Seabra, A. G. (2013). “Teste contrastivo de compreensão auditiva e de leitura,” in *Avaliação Neuropsicológica Cognitiva: Leitura, Escrita e Aritmética*, eds A. G. Seabra, N. M. Dias, and F. C. Capovilla (São Paulo: Memnon Edições Científicas), 3, 29–53
- Dias, N. M., Montiel, J. M., and Seabra, A. G. (2015). Development and interactions among academic performance, word recognition, listening, and reading comprehension. *Psicol. Reflex. Crit.* 28, 404–415. doi: 10.1590/1678-7153.201528221
- Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104
- Johnston, T. C., and Kirby, J. R. (2006). The contribution of naming speed to the simple view of reading. *Read. Writ.* 19, 339–361. doi: 10.1007/s11145-005-4644-2
- Joshi, M., Tao, S., Aaron, P. G., and Quiroz, B. (2012). Cognitive component of componential model of reading applied to different orthographies. *J. Learn. Disabil.* 45, 480–486. doi: 10.1177/0022219411432690
- Joshi, R. M., and Aaron, P. G. (2000). The component model of reading: simple view of reading made a little more complex. *Read. Psychol.* 21, 85–97. doi: 10.1080/02702710050084428
- Hudson, R., Lane, H., and Pullen, P. (2005). Reading fluency assessment and instruction: What, why, and how? *Read. Teach.* 58, 702–714. doi: 10.1598/rt.58.8.1
- Klauda, S. L., and Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *J. Educ. Psychol.* 100, 310–321. doi: 10.1037/0022-0663.100.2.310
- Kuhn, M. R., Schwanenflugel, P. J., and Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: automaticity, prosody, and definitions of fluency. *Read. Res. Q.* 45, 230–251. doi: 10.1598/RRQ.45.2.4
- Martins, M. A., and Capellini, S. A. (2014). Fluency and reading comprehension in students from 3rd to 5th of elementary school grades. *Estud. Psicol.* 31, 499–506.
- Montiel, J. M. (2008). *Validity Evidences of Tests for Identify Written Language Components*. Ph.D. thesis, Universidade São Francisco, Itatiba.
- National Reading Panel (US) and National Institute of Child Health and Human Development (US) (2000). *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction*. Washington, DC: National Institute of Child Health and Human Development, National Institutes of Health.
- Puliezi, S., and Maluf, M. R. (2014). Fluency and its importance to reading comprehension. *Psico USF* 19, 467–475. doi: 10.1590/1413-82712014019003009
- Seabra, A. G., and Capovilla, F. C. (2010). *Teste de Competência de Leitura de Palavras e Pseudopalavras*. São Paulo: Memnon.
- Turkylmaz, M., Can, R., Yildirim, K., and Ateş, S. (2014). Relations among Oral Reading Fluency, Silent Reading Fluency, Retell Fluency, and Reading Comprehension. *Proc. Soc. Behav. Sci.* 116, 4030–4034. doi: 10.1016/j.sbspro.2014.01.885
- Yovanoff, P., Duesbery, L., Alonzo, J. and Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educ. Meas. Issues Pract.* 24, 4–12. doi: 10.1111/j.1745-3992.2005.00014.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Seabra, Dias, Mecca and Macedo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Reading Fluency As a Predictor of School Outcomes across Grades 4–9

Lucia Bigozzi, Christian Tarchi\*, Linda Vagnoli, Elena Valente and Giuliana Pinto

Department of Education and Psychology, University of Florence, Florence, Italy

This study analyzed the predictive relationship between reading fluency and school outcomes across school levels (primary, secondary, and high school), after controlling on the effect of reading comprehension. The sample included 489 children attending Italian primary (grades 4 and 5), secondary (grades 6 and 8), and high schools (grade 9). Students' reading fluency and comprehension were examined with a standardized reading achievement test. At the end of the school year, we requested the school reports of each participant. According to our data, reading fluency predicted all school marks in all literacy-based subjects, with reading rapidity being the most important predictor. School level did not moderate the relationship between reading fluency and school outcomes, confirming the importance of effortless and automatized reading even in higher school levels. Overall this study emphasizes the importance of identifying evidence-based tasks that can be administered in a short time and to many different individuals, which are easy to create, and are linked to school outcomes.

**Keywords:** reading fluency, reading comprehension, school outcomes, school grades, predictors

## OPEN ACCESS

### Edited by:

Simone Aparecida Capellini,  
Sao Paulo State University, Brazil

### Reviewed by:

Kelly B. Cartwright,  
Christopher Newport University, USA  
Angela Jocelyn Fawcett,  
Swansea University, UK

### \*Correspondence:

Christian Tarchi  
christian.tarchi@unifi.it

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 11 April 2016

**Accepted:** 31 January 2017

**Published:** 14 February 2017

### Citation:

Bigozzi L, Tarchi C, Vagnoli L,  
Valente E and Pinto G (2017) Reading  
Fluency As a Predictor of School  
Outcomes across Grades 4–9.  
Front. Psychol. 8:200.  
doi: 10.3389/fpsyg.2017.00200

## INTRODUCTION

Teaching children to read fluently and comprehend a text is one of the main goals of early childhood education, because of the primary aims of reading which are to achieve one's goals, develop one's knowledge and potential, and participate in society (OECD, 2013). Reading is also a fundamental skill for school achievement (Hulme and Snowling, 2011), as also shown by studies documenting the persistence of reading disorders across the life span (Shaywitz et al., 1999). Reading fluency and comprehension are strictly inter-related, and also correlated with important aspects of academic life, such as school outcomes (Álvarez-Cañizo et al., 2015), or training success (Krumm et al., 2008). After primary school, teachers tend to focus on reading comprehension, neglecting the fostering of students' reading fluency, the influence of which is believed to fade on school outcomes. However, this assumption has recently been challenged, and the importance of reading fluency in adolescence re-evaluated (Rasinski et al., 2009; Ricketts et al., 2014; Zoccolotti et al., 2014). Moreover, recent literacy theories have documented how text use differs as a function of domains of academic subjects (Moje Birr et al., 2010), and reading strategies might become less generalizable as students move into increasingly specific disciplinary knowledge areas in higher school levels (Shanahan and Shanahan, 2008). This study analyzed the independent contribution of reading fluency to predict school outcomes in several subjects, after the effect of reading comprehension having been controlled for at three school levels, primary, secondary and high school in Italian students.



## The Contribution of Fluency to Reading

Reading fluency is defined as the ability to read rapidly, accurately, and with the proper expression, and includes three main components, reading rapidity, accuracy, and prosody (Kuhn and Stahl, 2003; Álvarez-Cañizo et al., 2015; Elhassan et al., 2015). Although all three components play an important role for school achievement, the first two ones (i.e., rapidity and accuracy) are most commonly assessed, in both educational and clinical contexts. As an implication, the only standardized measures available, at least in the Italian context, are reading rapidity and accuracy tests. Hence, this study will focus on the effect of reading fluency and accuracy on school achievement.

Reading fluency represents an extremely complex process, as the reader has to integrate perceptual skills to automatically translate letters into coherent sound representations, lexical skills to unitize those sound components into recognizable wholes, and processing skills to identify meaningful connections within and between sentences, relate text information with prior knowledge, and make inferences to fill in the gaps in the text (Fuchs et al., 2001). These skills need to be coordinated in a seemingly effortless manner: reading fluency reflects this complex integration and can be used as a reliable measure of reading expertise. Indeed, effortless and efficient reading fluency frees up cognitive resources for the higher-level and demanding comprehension processing of the text (Fuchs et al., 2001). In this regard, the theoretical foundation is represented by seminal article LaBerge and Samuel's (1974). According to these authors, we are able to perform two or more activities (e.g., fluency and comprehension) if we alternatively direct our attention between the activities, or if one of the activities is so mastered that it is performed automatically. In non-fluent readers, attention is drained by the decoding activity, and cognitive resources are not available for the comprehending activity (Pikulski and Chard, 2005).

Certainly, reading, decoding and comprehension components are correlated (Pinnell et al., 1995; Pikulski and Chard, 2005). But decoding and comprehension can be separated, or at least dissociated. For example, when diagnosing a learning disorder with an impairment in reading, the DSM-V requires us to specify whether word reading accuracy, reading rate or fluency, spelling, or reading comprehension are compromised (American Psychiatric Association, 2013). The possibility that a compromised fluency might be dissociated from a compromised comprehension represents evidence that the two skills are lodged in distinct mechanisms. The relationship between reading fluency and comprehension is complex, and it is difficult to determine whether the former is a cause or a consequence of the latter, although several studies suggest that reading fluency influences the reading comprehension process (Fuchs et al., 2001; Rasinski et al., 2005; Nese et al., 2013).

## Reading Fluency and School Outcomes

Regarding these two skills, reading fluency has been neglected by studies, especially in later school levels. However, reading fluency is an important topic area of longstanding interest, and it is currently receiving considerable attention (Tichá et al., 2009; Ari, 2015; Elhassan et al., 2015). Reading fluency represents

a crucial point for teachers to help struggling students meet school standards. Reading fluency becomes even more important in school settings based on learning from textbooks and time-limited assessment to determine students' outcomes. Indeed, many scholars consider reading fluency as a curriculum-based measurement, that is a valid and reliable procedure to monitor students' progress on a frequent basis and make instructional decisions (Tichá et al., 2009; Nese et al., 2013).

However, there is still disagreement on whether the magnitude of the correlation between reading fluency and high-stakes assessment scores declines across years (Reschly et al., 2009), or, instead, whether reading fluency is a key for successful school achievement even beyond elementary grades (Rasinski et al., 2005, 2009). This debate is still far from being settled, because the few studies conducted on this topic have led to contrasting results, neglecting the differences existing between languages (i.e., depth of orthography) and school system (i.e., school grades to measure students' outcomes). Regarding the first point, research on reading fluency growth is limited beyond grade 5 (Nese et al., 2013), and even those that explored the predictive power of reading fluency on school outcomes across grades lead to contradictory results (see Reschly et al.'s meta-analysis, 2009). Regarding the second point, learning to read in deep orthographies, characterized by an irregular mapping between letters and phonemes, is a much slower process than what happens in shallow orthographies, characterized by a regular mapping between letters and phonemes (Zoccolotti et al., 2008). In languages with a shallow orthography, such as Italian, reading accuracy is reached quite rapidly, making this parameter a less important indicator of reading proficiency and school outcomes than reading fluency (Bigozzi et al., 2016a,b). Thus, when assessing reading fluency, it is important to clearly distinguish the contribution of accuracy from that of rapidity. Regarding the second point, several studies have assessed students' school outcomes through standardized reading achievement tests. However, more recently some scholars have proposed school grades as a more ecological measure of school outcome. School grades have been criticized for low objectivity, reliability and validity (Krumm et al., 2008), but these issues have been criticized and disconfirmed by several studies that reported high correlations between school grades and other academic criteria (achievement tests, training success, and the like; Krumm et al., 2008; Rockoff and Speroni, 2010). A few authors even claimed that school grades might be a better predictor of graduation rates than standardized test scores, such as SAT scores (Bowen et al., 2009). School marks represent relevant real-life criteria to assess school outcomes (Krumm et al., 2008), and this is particularly true in countries in which school marks assigned by teachers represent the standard measure for students' school achievements and where students progress through school grades only if they have achieved at least a satisfactory level in each subject taught.

Finally, past research has demonstrated that comprehending and learning from text are associated but not overlapping processes (McNamara et al., 1996). Students might be able to achieve immediate comprehension of a text, but might not have learnt the concepts included in it. When students

become metacognitively aware of the importance of reading comprehension for school, they begin to put more effort into this process. Non-fluent readers might invest most of their cognitive resources in comprehending a text, and this task might drain cognitive capacity from studying the text for school achievement. In this sense, reading comprehension might have a positive effect on school outcomes if reading fluency is effective, efficient and effortless. To the best of our knowledge, prior studies have not tested whether reading fluency mediates the relationship between reading comprehension and school outcomes or not.

## Aims of the Study

The aim of this study was to analyze the predictive relationship between reading fluency and school outcomes across school levels (primary, secondary, and high school). More specifically, we expected that: (a) reading fluency contributes to predicting school marks in all school subjects in which reading plays a main role, after the effect of reading comprehension being controlled for; (b) reading fluency mediates the relationship between reading comprehension and school grades; (c) the contribution of reading fluency to school outcomes is not moderated by school level.

## METHODS

### Participants

The sample included 489 children attending Italian primary (grades 4 and 5), secondary (grades 6 and 8), and high schools (grade 9) in a mid-sized city in Central Italy (see **Table 1**). From this sample we had previously excluded foreign children and those who were covered by a certificate attesting the presence of a Learning Disability. The parents of the participants gave informed consent for the participation of their children in the study. The measures were administered at a time agreed on with the school and with due adherence to the requirements of privacy and informed consent required by Italian law (Legislative Decree DL-196/2003). Regarding the ethical standards for research, the study referred to the last version of the Declaration of Helsinki (World Medical Association, 2013). The present study was approved by the Ethics Committee of the Department of Psychology at the University of Florence, Italy.

In the Italian educational system, schools' programs are defined by the National Guidelines for the Curriculum, set by the Ministry of Education and Research. Students enter primary school at 6 years of age and stay for 5 years. Students enter secondary school at 11 years of age and this lasts for 3 years. At 14 years of age, students have to choose a specialization and enter high school, which lasts for 5 years. Class sizes are about 20 students in rural areas and small towns, and 30–35 students in large cities. The purpose of primary school is to teach the fundamental knowledge and skills to develop basic cultural competence. The subjects taught are: Italian, English, History, Geography, Mathematics, Science, Technology, Music, Art, and Physical Education. The timetable offers the following options: 24 h a week; 27 h a week; up to 30 h a week, involving up to 3 h per week for extra-curricular activities; or 40 h a week, including the lunchtime meal. In secondary school, the

**TABLE 1 | Sample description.**

School	Grade	N	Males	Females
Primary	4	143	77	66
	5	145	76	69
Secondary	6	70	34	36
	8	71	36	35
High	9	60	15	45
Total		489	238	251

minimum teaching time is 30 h per week. The subjects taught are: Italian (9 h per week), in-depth studies in literary subjects (1 h per week), Mathematics and Science (6 h per week), Technology (2 h per week), English (3 h per week), second foreign language (2 h per week), Art (2 h per week), Physical Education (2 h per week) and Music (2 h per week). In high school the timetable is 27 h per week. The subjects taught depend on the specialization of the school, but all schools include: Italian, English, History, Geography, Mathematics, Science, Technology, Music, Art, and Physical Education. Period assessments take place twice every year, at the end of each 4-month term. The evaluations in each subject are the responsibility of the teacher and are expressed in numerical marks out of 10 (from 0 to 10)<sup>1</sup>.

### Measures

Students' reading fluency and comprehension were examined with a standardized reading achievement test (MT Reading Test, Cornoldi and Colpo, 1995, 2011; Cornoldi et al., 2010). These tests are standardized instruments currently used in Italy for the assessment of reading processes (fluency and/or comprehension). Their reliability and validity has been well established in both, the construction of the instrument, and in several studies conducted by multiple investigators (e.g., Levorato et al., 2004; Faccioli et al., 2008; Angelelli et al., 2010; Zoccolotti et al., 2014). For a more accurate sample selection, we also considered the score on the standardized reading achievement test, and excluded students whose performance in reading accuracy, fluency, and/or comprehension was lower than the 5th percentile, following the indications of the DSM-5 (American Psychiatric Association, 2013).

### Reading Fluency

Each participant was tested individually by a trained experimenter. The participant was required to read the passage according to his or her grade level. Instructions emphasized accuracy and speed ("Read aloud as accurately and rapidly as you can.") while paying attention to the text content. The two components of reading fluency, rapidity (number of syllables read divided by time in seconds necessary to read them) and accuracy (number of words misread) were calculated. The following texts were assigned:

fourth grade, "L'indovina che indovino" ("The fortune-teller who guessed," 297 syllables);

<sup>1</sup>www.indire.it/lucabas/lkmw\_img/eurydice/quaderno\_eurydice\_30\_per\_web.pdf

fifth grade, “Vecchi proverbi” (“*Old sayings*,” 448 syllables); sixth grade, “Sogni a Hiroshima” (“*Dreams in Hiroshima*,” 592 syllables); eighth grade, “Città da salvare” (“*Cities to save*,” 576 syllables); ninth grade, “26 Dicembre 2004” (“*26th December 2004*,” 1123 syllables).

## Reading Comprehension

The reading comprehension test was collectively administered by a trained experimenter. The participant had to silently read a text and answer multiple-choice questions, with the possibility of accessing the text. Texts, number of questions (10 or 15) varied with school levels. Raw scores were converted to z scores according to standard reference data (Cornoldi and Colpo, 1995, 2011; Cornoldi et al., 2010). The following texts were assigned:

fourth grade, “Il leone e la leonessa” (“The lion and the lioness,” 10 questions, e.g., “What is the savannah?” 241 words); fifth grade, “Il viaggio delle anguille” (“The eels’ journey,” 10 questions, e.g., “Which ocean do eels cross?” 267 words); sixth grade, “Il pescatore, la volpe e l’orso” (“The fisherman, the fox, and the bear,” 15 questions, e.g., “Which of these sentences is more important for the development of the plot?” 503 words); eighth grade, “Don Orione” (“Father Orione,” 15 questions, e.g., “Which word would you choose to substitute the term ‘herd’?” 193 words); ninth grade, “Piaggia” (“*Piaggia*,” 15 questions, e.g., “Why was the rifle considered to be a prodigious weapon?” 333 words).

## School Marks

At the end of the school year, we requested the school reports of each participant, and took note of their scores in each subject (Italian, English as a foreign language, History, Geography, Mathematics, Sciences, Technology, Music, Art, and Physical Education). For most of the subjects we were able to collect data for all participants. As regards Technology, this subject is not taught in primary school, thus we collected data only for secondary and high school students, for a total of 234.

## Data Analysis

Each variable’s extreme outliers were identified and eliminated by observing the relative box-plots. Through examination of the skewness and kurtosis of each dependent variable’s probability distribution we verified that all variables were normally distributed, except for reading accuracy. Thus, reading accuracy was normalized through a monotonic transformation.

The first hypothesis (independent contribution of reading fluency on school marks) were explored through a hierarchical multiple linear regression analysis, with reading comprehension included in the first step, reading fluency in the second step, and school marks as dependent variables.

The second hypothesis (mediational effect of reading fluency on the association between reading comprehension and school grades) was explored through a mediation analysis. The third hypothesis (moderating role of school level on the association between reading fluency and school marks) was explored through a moderation analysis. Both mediation and moderation

analyses were conducted through PROCESS (release 2.15), an SPSS Macro created by Hayes (2012). The direct, indirect and moderation effects were derived from linear regression models. As suggested by Preacher and Hayes (2008), we used the bootstrapping strategy to test the mediation hypothesis, as it is the most powerful method to obtain confidence limits for specific indirect effects under most conditions. We used 5,000 bootstrap samples to construct bootstrap confidence intervals for indirect effect. The end points of bootstrap confidence intervals of the indirect effect were determined through the bias corrected method. In both tests, reading comprehension was included as a covariate. In the moderation analysis, the school moderating effect for each reading fluency component was tested by controlling the effect of the other component: when reading rapidity was the independent variable, we included reading accuracy as a covariate, and when reading accuracy was the independent variable, we included reading rapidity as a covariate.

## RESULTS

Descriptive statistics for all the variables included in the study are reported in **Table 2**.

The correlational analyses show that overall reading rapidity correlates positively with accuracy. The two reading fluency components are differently associated to reading comprehension. Reading rapidity and comprehension did not correlate at a statistically significant level, whereas reading accuracy and comprehension were negatively correlated: the fewer decoding mistakes made, the better the comprehension. Whereas reading rapidity and accuracy, and reading accuracy and comprehension correlated at all school levels, reading rapidity and comprehension correlated in grades 4, 5, 6, and 9, but not in grade 8 (see **Table 3**). **Table 4** shows the correlation between reading measures and school outcomes for the total sample, as well as for each grade.

The multiple regression analyses with hierarchical method showed that reading comprehension, inserted in step 1, significantly predicted school marks in all subjects ( $\Delta R^2$  ranged from .03 in Music to .16 in History), whereas reading fluency, inserted in step 2, predicted only grades in Italian, English as a foreign language, History, Geography, Mathematics, and Sciences, with  $\Delta R^2$  ranging between .03 and .06. In terms of specific contributions, Italian, English as a foreign language, and History were predicted by both components of reading fluency, Geography and Sciences were predicted by reading rapidity only, and Mathematics by reading accuracy only (see **Table 5**).

The contribution of reading comprehension to school outcomes was mostly direct, as the mediational effect of reading rapidity and/or accuracy were either non-significant (for Geography, Sciences, and Technology) or significant but marginal (for Italian, English, History, Mathematics, Music, Art, and Physical Education). Variances explained by partially mediated models ranged between 3 and 14% (see **Table 6**).

Overall, school moderated the relationship between reading fluency and school outcomes, although the percentages of variance explained by the interaction between school and reading

**TABLE 2 | Descriptive statistics of all variables: participants, mean, minimum, maximum, standard deviation, skewness and kurtosis.**

	<i>N</i>	Min	Max	<i>M</i> ± <i>SD</i>	Skewness	Kurtosis
Reading fluency – rapidity (syllables/seconds)	489	1.21	7.48	4.25 ± 1.18	0.22 ± 0.11	−0.32 ± 0.22
Reading fluency – accuracy	487	0	18	2.12 ± 2.29	2.82 ± 0.11	12.49 ± 0.22
Reading comprehension	488	2	15	9.04 ± 0.18	−0.33 ± 0.11	0.59 ± 0.22
Italian	489	6	10	8.04 ± 1.14	−0.04 ± 0.11	−0.74 ± 0.22
English as a foreign language	489	6	10	8.10 ± 1.18	−0.13 ± 0.11	−0.81 ± 0.22
History	489	6	10	8.10 ± 1.24	−0.20 ± 0.11	−0.96 ± 0.22
Geography	489	6	10	8.02 ± 1.21	−0.17 ± 0.11	−0.92 ± 0.22
Mathematics	489	6	10	7.95 ± 1.20	−0.02 ± 0.11	−0.87 ± 0.22
Sciences	489	5	10	8.14 ± 1.18	−0.27 ± 0.11	−0.78 ± 0.22
Technology	387	6	10	8.19 ± 1.05	−0.11 ± 0.12	−0.59 ± 0.25
Music	489	6	10	8.36 ± 0.97	−0.19 ± 0.11	−0.20 ± 0.22
Art	489	6	10	8.50 ± 0.98	−0.25 ± 0.11	−0.30 ± 0.22
Physical Education	489	6	10	8.65 ± 0.91	−0.29 ± 0.11	−0.08 ± 0.22

**TABLE 3 | Correlational analysis of reading measures: rapidity, accuracy, and comprehension.**

	Total	Grade 4	Grade 5	Grade 6	Grade 8	Grade 9
Rapidity*Accuracy	−0.51**	−0.43**	−0.49**	−0.61**	−0.48**	−0.50**
Rapidity*Comprehension	0.05	0.43**	0.30**	0.38**	0.18	0.41**
Accuracy*Comprehension	−0.17**	−0.44**	−0.29**	−0.34**	−0.37**	−0.42**

\* $p < 0.05$ , \*\* $p < 0.01$ .

**TABLE 4 | Correlations between reading measures (rapidity, accuracy and comprehension) and school outcomes (Italian, English as a foreign language, History, Geography, Mathematics, Sciences, Technology, Music, Art, and Physical education) for the total sample ( $n = 489$ ), and broken down by grade: 4 ( $n = 143$ ), 5 ( $n = 145$ ), 6 ( $n = 70$ ), 8 ( $n = 71$ ), and 9 ( $n = 60$ ).**

		Ita	Eng	His	Geo	Mat	Sci	Tech	Mus	Art	P.E.
Tot	Rap.	0.23**	0.24**	0.20**	0.22**	0.17**	0.20**	0.07	0.08	0.01	−0.05
	Acc.	−0.25**	−0.26**	−0.23**	−0.21**	−0.25**	−0.20**	−0.08	−0.14**	−0.09*	−0.05
	Comp.	0.38**	0.37**	0.41**	0.34**	0.31**	0.41**	0.27**	0.17**	0.24**	0.21**
Gr. 4	Rap.	0.40**	0.48**	0.44**	0.39**	0.35**	0.38**	0.26**	0.23**	0.19*	0.18*
	Acc.	−0.30**	−0.36**	−0.32**	−0.26**	−0.30**	−0.25**	−0.14	−0.24**	−0.24**	−0.15
	Comp.	0.55**	0.56**	0.61**	0.56**	0.51**	0.58**	0.42**	0.34**	0.33**	0.30**
Gr. 5	Rap.	0.47**	0.31**	0.43**	0.50**	0.43**	0.52**	0.38**	0.28**	0.17*	0.09
	Acc.	−0.33**	−0.24**	−0.32**	−0.32**	−0.31**	−0.34**	−0.22*	−0.18*	−0.22**	−0.15
	Comp.	0.32**	0.25**	0.33**	0.29**	0.35**	0.35**	0.25**	0.23**	0.19*	0.19*
Gr. 6	Rap.	0.61**	0.59**	0.42**	0.41**	0.38**	0.45**	0.30	0.18	0.47**	0.21
	Acc.	−0.44**	−0.45**	−0.43**	−0.38**	−0.44**	−0.36**	−0.21	−0.14	−0.21	−0.04
	Comp.	0.67**	0.61**	0.72**	0.63**	0.59**	0.63**	0.41*	0.26*	0.27*	0.28*
Gr. 8	Rap.	0.50**	0.57**	0.51**	0.53**	0.36**	0.45**	0.47**	0.58**	0.33**	−0.04
	Acc.	−0.23	−0.40**	−0.35**	−0.35**	−0.29*	−0.32**	−0.27*	−0.20	−0.09	−0.13
	Comp.	0.38**	0.28*	0.36**	0.37**	0.34**	0.25*	0.31**	0.12	0.11	0.08
Gr. 9	Rap.	0.78**	0.71**	0.63**	0.59**	0.40**	0.42**	0.41**	0.48**	0.44**	0.22
	Acc.	−0.55**	−0.51**	−0.39**	−0.38**	−0.33*	−0.32*	−0.39**	−0.42**	−0.37**	−0.35**
	Comp.	0.54**	0.53**	0.48**	0.50**	0.38**	0.43**	0.46**	0.50**	0.49**	0.41**

\*\* $p < 0.01$ ; \* $p < 0.05$ . Gr., Grade; Tot., Total; Rap., Rapidity; Acc., Accuracy; Comp., Comprehension; Ita, Italian; Eng, English as a foreign language; His, History; Geo, Geography; Mat, Mathematics; Sci, Sciences; Tech, Technology; Mus, Music; P.E., Physical Education.

fluency were small (between 1 and 5%). For Italian, reading rapidity contributed to explaining variance in students' outcome in primary and high school, but not in secondary school. The

analysis of confidence intervals, instead, did not confirm the moderation effect for reading accuracy. For English, reading rapidity contributed to explaining variance in students' outcome



**TABLE 5 | Results from the multiple regression analysis with hierarchical method to control for the effect of reading fluency on school grades.**

	Step 1: RC	Step 2: RF		Total R <sup>2</sup>
	$\Delta R^2$	$\Delta R^2$	Rapidity ( $\beta$ ) Accuracy ( $\beta$ )	
Italian	0.14**	0.05**	0.14** −0.12*	0.19**
English as a foreign language	0.13**	0.06**	0.14** −0.13*	0.18**
History	0.16**	0.04*	0.12* −0.11*	0.20**
Geography	0.11**	0.04*	0.15** −0.08	0.15**
Mathematics	0.09**	0.04**	0.05 −0.18**	0.13**
Sciences	0.16**	0.03**	0.13** −0.07	0.19**
Technology	0.07**	0.00	0.04 −0.01	0.06**
Music	0.03**	0.01	−0.00 −0.11*	0.03**
Art	0.05**	0.01	−0.07 −0.10	0.05**
Physical Education	0.04**	0.01	−0.12* −0.09	0.05*

\* $p < 0.05$ ; \*\* $p < 0.01$ . RC, Reading comprehension; RF, Reading Fluency; ns, non-significant.

in primary and high school, but not in secondary school; reading accuracy contributed to explaining variance in students' outcome only in high school, but not in primary and secondary school. For history, reading rapidity contributed to explaining variance in students' outcome in primary and high school, but not in secondary school; reading accuracy contributed to explaining variance in students' outcome only in primary school, but not in primary or secondary school. For geography, reading rapidity contributed to explaining variance in students' outcome in primary and high school, but not in secondary school. For mathematics, school did not moderate the relationship between reading fluency and students' outcome. For sciences, reading rapidity contributed to explaining variance in students' outcome in primary school, but not in secondary and high school. For music, reading accuracy contributed to explaining variance in students' outcome in high school, but not in primary or secondary school. For physical education, reading rapidity contributed to explaining variance in students' outcome in secondary school, but in primary or high school (see **Table 7**).

## DISCUSSION

The aim of this study was to analyze whether reading fluency influences students' school outcomes in school subjects, independently of the effect of reading comprehension, and whether the independent contribution of reading fluency is moderated by the school level. According to the correlational analysis, reading fluency and comprehension are associated, confirming prior studies (Pinnell et al., 1995; Pikulski and Chard, 2005). More specifically, reading accuracy appears to be more strictly associated with reading comprehension than reading rapidity, confirming that reading "fast" does not help children to adequately process the information included in the text. It is important to address a deep form of reading fluency, according to which this construct is part of a developmental

process of building decoding skills that are reciprocally and causally connected with reading comprehension, rather than just be considered as "fast reading" (Pikulski and Chard, 2005).

This study emphasizes the fundamental importance of reading comprehension and fluency in students' school outcomes. Reading comprehension and fluency are strictly inter-related processes, however, according to our data, both contribute independently to school marks in several subjects.

Firstly, although several studies suggest that reading fluency influences the reading comprehension process (Fuchs et al., 2001; Rasinski et al., 2005; Nese et al., 2013), results from the mediational analyses showed that the contribution of reading comprehension to school marks is mostly direct. More importantly, this study contributes to re-evaluating the role played by reading fluency, and confirms that effortless and automatic reading fluency frees up important cognitive resources for the comprehension activity, a high-level and demanding process (Fuchs et al., 2001; Pikulski and Chard, 2005; Tichá et al., 2009; Nese et al., 2013). The efficacy of reading fluency is especially significant for subjects in which literacy skills and textbook studying play a primary role (i.e., Italian, English, History, Geography, Mathematics, and Sciences). Reading rapidity was the most important predictor among the two reading fluency components: as suggested by several scholars, in shallow orthographies, reading accuracy is reached rapidly, which makes reading rapidity a much more important indicator of reading proficiency (Zoccolotti et al., 2008; Pinto et al., 2015). Instead, reading accuracy played an important role for Geography and Mathematics, probably as these subjects involve more focused attention on visuo-spatial elements, besides the verbal one (Schnotz, 2002). Many textbooks require students to mostly process verbal information, whereas in Geography and Mathematics, students need to integrate verbal and graphic information, an activity that requires a slower and more accurate processing (Massey and Riley, 2013).

The moderation analysis contributed to our understanding of the relationship between reading fluency and school outcomes. Overall, the effect was mainly confirmed for primary school, when students are in the process of reading acquisition (Pinto et al., 2015). In secondary school instead, reading fluency appears to be neglected, except for Italian in which there is still a strong emphasis on grammar. In secondary school reading fluency did not influence students' outcomes. The importance of reading fluency for school outcomes in primary school is not questionable, since the teachers' focus at this level is on basic literacy skills (Firestone and Herriott, 1982; Alvermann and Moore, 1991). Once in secondary school, the focus shifts to subject-matter literacy (Knott, 1986; Alvermann and Moore, 1991). However, secondary school instruction mainly puts emphasis on factual textual information, with textbooks acting as sources of information (Smith and Feathers, 1983; Alvermann and Moore, 1991). Consequently, students do not need to dedicate many cognitive resources to the reading comprehension process, which also allows poor decoders to achieve good school outcomes. Instead, at high school level, reading fluency brings again an independent contribution to school outcomes. Several reasons are able to explain this result: (i) in high school

**TABLE 6 | Results from the mediation analysis to control for the mediational effect of reading fluency (rapidity and accuracy) on the association between reading comprehension and school grades.**

	Model	Total effect	Direct effect	Indirect effect of rapidity			Indirect effect of accuracy		
	$R^2$	$\beta$	$\beta$	$\beta$	BootLLCI	BootULCI	$\beta$	BootLLCI	BootULCI
Italian	0.14**	0.20**	0.19**	0.004	−0.002	0.014	0.008*	0.001	0.021
English	0.13**	0.19**	0.18**	0.004	−0.001	0.014	0.009*	0.002	0.024
History	0.16**	0.23**	0.22**	0.003	−0.002	0.014	0.008*	0.001	0.022
Geography	0.11**	0.18**	0.17**	0.004	−0.003	0.016	0.006	−0.001	0.018
Mathematics	0.09**	0.17**	0.15**	0.001	−0.001	0.010	0.013*	0.004	0.028
Sciences	0.16**	0.22**	0.21**	0.004	−0.002	0.015	0.005	−0.001	0.017
Technology	0.07**	0.12**	0.12*	0.001	−0.001	0.009	0.001	−0.008	0.011
Music	0.03**	0.07**	0.07**	0.002	−0.005	0.003	0.004*	0.001	0.018
Art	0.05**	0.10**	0.10**	−0.002**	−0.009	0.001	0.006*	0.000	0.017
Physical Education	0.04**	0.08**	0.08**	−0.003	−0.011	0.001	0.005*	0.000	0.015

\* $p < 0.05$ ; \*\* $p < 0.01$ . ns, non significant.

**TABLE 7 | Results from the moderation analysis to control for the effect of school level on the association between reading fluency and school grades, with reading comprehension as a covariate, reading accuracy as a covariate for reading rapidity, and reading rapidity as a covariate for reading accuracy ( $R^2$ ,  $R^2$  change, unstandardized regression coefficients, lower and higher confidence intervals).**

DV	IV	Model	Interaction	Primary			Secondary			High		
		$R^2$	$\Delta R^2$	$\beta$	LLCI	UCLI	$\beta$	LLCI	UCLI	$\beta$	LLCI	UCLI
Italian	Rap.	0.36**	0.05**	0.25**	0.13	0.38	0.05	−0.10	0.18	1.13**	0.82	10.45
	Acc.	0.34**	0.03**	−0.46*	−0.92	−0.01	0.79*	0.17	1.40	−1.91**	−2.98	−0.84
English	Rap.	0.31**	0.03**	0.25**	0.12	0.38	0.09	−0.06	0.24	0.97**	0.63	1.31
	Acc.	0.29**	0.02**	−0.45	−0.94	0.04	0.34	−0.32	1.00	−1.70**	−2.86	−0.55
History	Rap.	0.33**	0.03**	0.31**	0.17	0.45	−0.02	−0.18	0.14	0.67**	0.32	1.02
	Acc.	0.32**	0.01**	−0.59*	−1.09	−0.08	0.64	−0.03	1.32	−0.61	−1.79	0.57
Geography	Rap.	0.33**	0.03**	0.37**	0.24	0.50	−0.02	−0.18	0.14	0.63**	0.28	0.97
Maths	Acc.	0.27**	0.00	−0.67**	−1.18	−0.16	−0.13	−0.82	0.55	−0.82	−2.02	0.36
Sciences	Rap.	0.27**	0.02**	0.33**	0.19	0.46	−0.04	−0.20	0.12	0.35	−0.01	0.70
Music	Acc.	0.21**	0.02**	−0.27	−0.69	0.16	0.28	−0.30	0.85	−1.45**	−2.45	−0.45
P.E.	Rap.	0.12**	0.03**	−0.02	−0.14	0.09	−0.34**	−0.47	−0.21	−0.01	−0.30	0.29

\*\* $p < 0.01$ , \* $p < 0.05$ . Rap., Rapidity; Acc., Accuracy; P. E., Physical Education.

basic literacy skills are no longer sufficient, because of the complexity of textbooks (Lester and Cheek, 1997); (ii) students who enter high school with a lack of sufficient reading fluency are not likely to find instructional support from teachers and/or remedial programs (Rasinski et al., 2005; Joseph and Schisler, 2009), and (iii) slow readers require significantly more time in accomplishing school tasks than normally-reading readers do, which might eventually lead to frustration, task-avoidance behaviors, and school failure (Rasinski et al., 2005; Archer et al., 2013). Overall, these results confirm the importance of reading fluency even in adolescence (Rasinski et al., 2009; Ricketts et al., 2014; Zoccolotti et al., 2014).

Overall this study emphasizes the importance of identifying evidence-based tasks that can be administered in a brief time and by many different individuals, which are easy to create, and are linked to school outcomes (Tichá et al., 2009; Nese et al., 2013). As the effect of reading fluency on school outcomes does not

fade after primary school, secondary and high school teachers should not underestimate the negative impact of ineffective and non-automatic reading fluency has on students' learning. This study also contributes to extend what we know about learning disorders on normally-developing children. In shallow orthographies, dyslexic readers have compromised accuracy and reading rapidity (Bigozzi et al., 2016a), and these compromised processes hinder student learning. This study confirms the same effect for the population of students without a learning disorder.

The results of this study are affected by a few limitations. A few intervening variables might explain the relationship between reading and learning, both higher-order (e.g., studying skills, metacognitive variables, or motivational variables, see Schiefele et al., 2012), and lower order ones (e.g., verbal ability, see Tilstra et al., 2009). Future studies should also include these variables in the research design to better explain under which conditions reading fluency fosters students' learning and

their school outcomes. Moreover, although cross-sectional data can be used to test mediation, longitudinal data are more appropriate. Future studies should replicate the results of this study through a longitudinal research design. Finally, few authors have emphasized the importance of reading prosody, besides rapidity and accuracy, for school achievement (Kuhn and Stahl, 2003). However, there is a lack of standardized measures of this reading fluency component, and future studies should aim at first validating reading prosody assessment and then analyzing the specific contribution of this component on school achievement.

Reading fluency is typically considered an important process for school achievement in beginning readers and in dyslexic students. This study provides new insights into the importance of fluent reading for academic outcomes beyond reading

comprehension. The shift from reading comprehension (more common in fluency research) to academic performance as the criterion variable in the study is novel and yielded important findings for the field. Our results contribute to renew the attention to specific processes (in this case, reading fluency) for school achievement, besides more general processes (such as intelligence), which can also be improved as a result of targeted interventions.

## AUTHOR CONTRIBUTIONS

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

## REFERENCES

- Álvarez-Cañizo, M., Suárez-Coalla, P., and Cuetos, F. (2015). The role of reading fluency in children's text comprehension. *Front. Psychol.* 6:1810. doi: 10.3389/fpsyg.2015.01810
- Alvermann, D. E., and Moore, D. W. (1991). "Secondary school reading," in *Handbook of Reading Research, Vol. II*, eds R. Barr, M. L. Kamil, P. B. Mosenthal, and P. D. Pearson (New York, NY: Longman), 951–983.
- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edn*. Washington, DC: American Psychiatric Association.
- Angelelli, P., Notarnicola, A., Judica, A., Zoccolotti, P., and Luzzatti, C. (2010). Spelling impairments in Italian dyslexic children: phenomenological changes in primary school. *Cortex* 46, 1299–1311. doi: 10.1016/j.cortex.2010.06.015
- Archer, A. L., Gleason, M. M., Vachon, V. L., Archer, A. L., Gleason, M. M., and Vachon, V. L. (2013). Decoding and fluency: foundation skills for struggling older readers. *Learn. Disabil. Q.* 26, 89–101. doi: 10.2307/1593592
- Ari, O. (2015). Fluency gains in struggling college readers from wide reading and repeated readings. *Read. Psychol.* 36, 270–297. doi: 10.1080/02702711.2013.864361
- Bigozzi, L., Tarchi, C., Pezzica, S., and Pinto, G. (2016a). Evaluating the predictive impact of an emergent literacy model on dyslexia in Italian children: A four-year prospective cohort study. *J. Learn. Disabil.* 49, 51–65. doi: 10.1177/0022219414522708
- Bigozzi, L., Tarchi, C., Pinto, G., and Accorti Gamannossi, B. (2016b). Predicting Dyslexia in a Transparent orthography from grade 1 literacy skills: a prospective cohort study. *Read. Writ. Q.* 32, 353–372. doi: 10.1080/10573569.2014.988310
- Bowen, W. G., Chingos, M. M., and McPherson, M. S. (2009). *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, NJ: Princeton University Press.
- Cornoldi, C., and Colpo, G. (1995). *Nuove Prove di Lettura MT per la Scuola Media Inferiore* (en. tr. Reading Test for Secondary School). Florence, IT: Giunti O.S.
- Cornoldi, C., and Colpo, G. (2011). *Prove di Lettura MT-2 per la Scuola Primaria* (en. tr. Reading Test for Primary School). Firenze, IT: Giunti O.S.
- Cornoldi, C., Pra Baldi, A., and Friso, G. (2010). *Prove MT Avanzate – 2 di Lettura per il Biennio Della Scuola Secondaria di II Grado* (en. tr. Reading Test for High School). Firenze, IT: Giunti O.S.
- Elhassan, Z., Crewther, S. G., Bavin, E. L., and Crewther, D. P. (2015). Preliminary validation of FastaReada as a measure of reading fluency. *Front. Psychol.* 6:1634. doi: 10.3389/fpsyg.2015.01634
- Faccioli, C., Peru, A., Rubini, E., and Tassinari, G. (2008). Poor readers but compelled to read: stroop effects in developmental dyslexia. *Child Neuropsychol.* 14, 277–283. doi: 10.1080/09297040701290040
- Firestone, W. A., and Herriott, R. E. (1982). Prescriptions for effective elementary schools don't fit secondary schools. *Educ. Leadersh.* 40, 51–53.
- Fuchs, L. S., Fuchs, D., Hosp, M. K., and Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: a theoretical, empirical, and historical analysis. *Sci. Stud. Read.* 5, 239–256. doi: 10.1207/S1532799XSSR0503\_3
- Hayes, A. F. (2012). *PROCESS: A Versatile Computational Tool for Observed Variable Mediation, Moderation, and Conditional Process Modeling [White paper]*. Available online at: <http://www.afhayes.com/>
- Hulme, C., and Snowling, M. J. (2011). Children's reading comprehension difficulties: nature, causes, and treatments. *Curr. Direct. Psychol. Sci.* 20, 139–142. doi: 10.1177/0963721411408673
- Joseph, L. M., and Schisler, R. (2009). Should adolescents go back to the basics? A review of teaching word reading skills to middle and high school students. *Remedial Spec. Educ.* 30, 131–148. doi: 10.1177/0741932508315646
- Knott, G. (1986). Secondary school contexts of reading and writing instructions. *Theory Pract.* 25, 77–83. doi: 10.1080/00405848609543204
- Krumm, S., Ziegler, M., and Buehner, M. (2008). Reasoning and working memory as predictors of school grades. *Learn. Individ. Differ.* 18, 248–257. doi: 10.1016/j.lindif.2007.08.002
- Kuhn, M. R., and Stahl, S. A. (2003). Fluency: a review of developmental and remedial practices. *J. Educ. Psychol.* 95, 3–21. doi: 10.1037/0022-0663.95.1.3
- LaBerge, D., and Samuel, S. J. (1974). Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. doi: 10.1016/0010-0285(74)90015-2
- Lester, J. H., and Cheek, E. H. (1997). The "real" experts address textbook. *J. Adolesc. Adult Lit.* 41, 282–291.
- Lavorato, M. C., Nesi, B., and Cacciari, C. (2004). Reading comprehension and understanding idiomatic expressions: A developmental study. *Brain Lang.* 91, 303–314. doi: 10.1016/j.bandl.2004.04.002
- Massey, D., and Riley, L. (2013). Reading math textbooks: An algebra teacher's patterns of thinking. *J. Adolesc. Adult Lit.* 56, 577–586. doi: 10.1002/JAAL.187
- McNamara, D. S., Kintsch, E., Songer, N. B., and Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cogn. Instr.* 14, 1–43. doi: 10.1207/s1532690xci1401\_1
- Moje Birr, E., Stockdill, D., Kim, K., and Kim, H. (2010). "The role of text in disciplinary learning," in *Handbook of Reading Research, Vol. 4*, eds M. L. Kamil, P. D. Pearson, E. Moje Birr, and P. P. Afflerbach (New York, NY: Taylor and Francis Group), 453–486.
- Nese, J. F. T., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., and Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across Grades 1–8. *J. Sch. Psychol.* 51, 625–642. doi: 10.1016/j.jsp.2013.05.006
- OECD (2013). *Pisa 2015 Draft Reading Literacy Framework March 2013*. Available online at: [http://www.oecd.org/pisa/pisaproducts/Draft\\_PISA\\_2015\\_Reading\\_Framework.pdf](http://www.oecd.org/pisa/pisaproducts/Draft_PISA_2015_Reading_Framework.pdf)
- Pikulski, J. J., and Chard, D. J. (2005). Fluency: bridge between decoding and reading comprehension. *Read. Teach.* 58, 510–519. doi: 10.1598/RT.58.6.2
- Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., and Beatty, A. A. (1995). *Listening to Children Read Aloud: Data from NAEP's Integrated Reading Performance Record (IRPR) at Grade 4*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

- Pinto, G., Bigozzi, L., Tarchi, C., Accorti Gamannossi, B., and Canneti, L. (2015). Cross-lag analysis of longitudinal associations between primary school students' writing and reading skills. *Read. Writ.* 28, 1–23. doi: 10.1007/s11145-015-9569-9
- Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40, 879–891. doi: 10.3758/BRM.40.3.879
- Rasinski, T., Padak, N. D., McKeon, C. A., Wilfong, L. G., Friedauer, J. A., and Heim, P. (2005). Is reading fluency a key for successful high school reading? *J. Adolesc. Adult Lit.* 49, 22–27. doi: 10.1598/JAAL.49.1.3
- Rasinski, T., Rikli, A., and Johnston, S. (2009). Reading fluency: more than automaticity? More than a concern for the primary grades? *Lit. Res. Instr.* 48, 350–361. doi: 10.1080/19388070802468715
- Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., and Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: a meta-analysis of the correlational evidence. *J. Sch. Psychol.* 47, 427–469. doi: 10.1016/j.jsp.2009.07.001
- Ricketts, J., Sperring, R., and Nation, K. (2014). Educational attainment in poor comprehenders. *Front. Psychol.* 5:445. doi: 10.3389/fpsyg.2014.00445
- Rockoff, J. E., and Speroni, C. (2010). Subjective and Objective Evaluations of Teacher Effectiveness. *Am. Econ. Rev.* 100, 261–266. doi: 10.1257/aer.100.2.261
- Schiefele, U., Schaffner, E., Möller, J., Wigfield, A., Nolen, S., and Baker, L. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Read. Res. Q.* 47, 427–463. doi: 10.1002/RRQ.030
- Schnotz, W. (2002). Commentary: towards an integrated view of learning from text and visual displays. *Educ. Psychol. Rev.* 14, 101–120. doi: 10.1023/A:1013136727916
- Shanahan, T., and Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: rethinking content-area literacy. *Harv. Educ. Rev.* 78, 40–60. doi: 10.17763/haer.78.1.v62444321p602101
- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., et al. (1999). Persistence of dyslexia: the connecticut longitudinal study at adolescence. *Pediatrics* 104, 1351–1359. doi: 10.1542/peds.104.6.1351
- Smith, F. R., and Feathers, K. M. (1983). Teacher and students perceptions of content area reading. *J. Read.* 26, 348–354.
- Tichá, R., Espin, C. A., and Wayman, M. M. (2009). Reading progress monitoring for secondary-school students: Reliability, validity, and sensitivity to growth of reading-aloud and maze-selection measures. *Learn. Disabil. Res. Pract.* 24, 132–142. doi: 10.1111/j.1540-5826.2009.00287.x
- Tilstra, J., McMaster, K., Van Den Broek, P., Kendeou, P., and Rapp, D. (2009). Simple but complex: components of the simple view of reading across grade levels. *J. Res. Read.* 32, 383–401. doi: 10.1111/j.1467-9817.2009.01401.x
- World Medical Association (2013). *Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects*. Fortaleza. Available online at: <http://www.wma.net/en/30publications/10policies/b3/index.html>
- Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., and Martelli, M. (2008). Reading development in an orthographically regular language: effects of length, frequency, lexicality and global processing ability. *Read. Writ.* 22, 1053–1079. doi: 10.1007/s11145-008-9144-8
- Zoccolotti, P., De Luca, M., Marinelli, C. V., and Spinelli, D. (2014). Modeling individual differences in text reading fluency: a different pattern of predictors for typically developing and dyslexic readers. *Front. Psychol.* 5:1374. doi: 10.3389/fpsyg.2014.01374

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bigozzi, Tarchi, Vagnoli, Valente and Pinto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Silent Reading Fluency and Comprehension in Bilingual Children

Beth A. O'Brien<sup>1\*</sup> and Sebastian Wallot<sup>2</sup>

<sup>1</sup> Education and Cognitive Development Lab, National Institute of Education, Nanyang Technological University, Singapore, Singapore, <sup>2</sup> Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

This paper focuses on reading fluency by bilingual primary school students, and the relation of text fluency to their reading comprehension. Group differences were examined in a cross-sectional design across the age range when fluency is posed to shift from word-level to text-level. One hundred five bilingual children from primary grades 3, 4, and 5 were assessed for English word reading and decoding fluency, phonological awareness, rapid symbol naming, and oral language proficiency with standardized measures. These skills were correlated with their silent reading fluency on a self-paced story reading task. Text fluency was quantified using non-linear analytic methods: recurrence quantification and fractal analyses. Findings indicate that more fluent text reading appeared by grade 4, similar to monolingual findings, and that different aspects of fluency characterized passage reading performance at different grade levels. Text fluency and oral language proficiency emerged as significant predictors of reading comprehension.

## OPEN ACCESS

### Edited by:

Giseli Donadon Germano,  
Universidade Estadual Paulista, Brazil

### Reviewed by:

Nicole D. Anderson,  
MacEwan University, Canada  
Angela Jocelyn Fawcett,  
Swansea University, UK

### \*Correspondence:

Beth A. O'Brien  
beth.obrien@nie.edu.sg

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 06 June 2016

**Accepted:** 09 August 2016

**Published:** 31 August 2016

### Citation:

O'Brien BA and Wallot S (2016) Silent  
Reading Fluency and Comprehension  
in Bilingual Children.  
Front. Psychol. 7:1265.  
doi: 10.3389/fpsyg.2016.01265

**Keywords:** text reading fluency, bilingual readers, silent reading, comprehension, recurrence quantification analysis, fractal analysis

## INTRODUCTION

An increasing proportion of individuals worldwide grow up bilingual or multilingual (Grosjean, 2013). Following this, many children learn to read for the first time in what would be a second language (McBride-Chang, 2004). Therefore, it is important to understand what contributes to reading proficiency and comprehension for bi- or multilingual individuals.

The simple view of reading framework (Hoover and Gough, 1990) places reading comprehension as a product of word reading (decoding) and listening comprehension. There is ample evidence from monolingual research to support this view, but more recent findings suggest a role for fluent reading of text which serves as a bridge between word decoding and reading comprehension (Pikulski and Chard, 2005; Adolf et al., 2006; Bashir and Hook, 2009). Reading fluency is found to mediate the relation between reading comprehension and decoding (Silverman et al., 2013), and also to partially mediate the relation between reading comprehension and listening comprehension (Kim and Wagner, 2015).

For second language learners, the simple view of reading is found to hold as it does for monolinguals (Verhoeven and van Leeuwe, 2012), but the role of reading fluency for bilingual or second language readers is not clear. While reading fluency is related to reading comprehension for both monolinguals (Fuchs et al., 2001; Hosp and Fuchs, 2005) as well as L2 readers (Baker and Good, 1995; DeRamirez and Shapiro, 2006), reading fluency research with bilingual children is scarce.

The role of text fluency to reading comprehension for bilingual children requires further investigation for several reasons. First, reading fluency is characterized by automatic word recognition (Samuels, 2006; Kuhn et al., 2010), and developing automaticity for word identification can be a challenge for bilingual readers (e.g., Van Heuven et al., 1998; Segalowitz and Hulstijn, 2005). Bilingual children and adults show slower lexical access (Gollan et al., 2005; Costa et al., 2006; Liu et al., 2010; Sandoval et al., 2010) and smaller receptive vocabularies (Oller et al., 2007; Bialystok et al., 2010) compared with monolingual peers. This may impact on their word fluency. Second, bilingual readers demonstrate poorer reading comprehension that persists despite adequate decoding skills (Proctor et al., 2005; Lesaux et al., 2008; Chen et al., 2012), findings which may be related to poorer text reading fluency found with such readers' performance (Geva et al., 1997; Crosson and Lesaux, 2010; Geva and Farnia, 2012).

Another issue to be resolved is how reading comprehension is related to text fluency *compared with* oral language proficiency in bilingual readers. In most studies, oral reading tasks are used to gauge text fluency, but oral reading may be more heavily influenced by oral language proficiency for bilingual individuals. For instance, lower levels of proficiency may result in slower articulation or mispronunciations during oral reading of a second or additional language. This may result in dysfluent reading (as measured with oral reading rate or prosody), but may not reflect the reader's comprehension of the text (e.g., Geva, 2006). Thus, oral reading fluency may not be an adequate gauge of reading proficiency for L2 or bilingual readers (Piper et al., 2016). For adult bilinguals the relation of oral reading fluency to reading comprehension is found to be weaker for L2 compared with L1 readers, ranging from correlations of 0.26 (Lems, 2012) to 0.46 and 0.51 for L2 adults (Jeon, 2012; Jiang et al., 2012), compared with a correlation of 0.80 for L1 children (Fuchs et al., 2001). With children, oral reading fluency may overestimate bilingual children's comprehension, if their word decoding skills are more advanced than their listening comprehension. Lems (2006) noted that when mastery of decoding skills precedes vocabulary development in L2 readers, they may decode without comprehension, as anecdotally noted by ELL teachers (DeRamirez and Shapiro, 2006). This point was supported by Jackson and Lu's (1992) report of a dissociation between English oral language scores and reading fluency in a group of precocious preschool L2 readers. That is, these children could read as fluently as native speakers did, but had significantly poorer oral language proficiency. Studies with primary school children also showed that oral language proficiency contributes additional variance to comprehension and that it has a moderating effect on the fluency-comprehension relation (Crosson and Lesaux, 2010; Geva and Farnia, 2012).

Thus, fluency measured with a silent reading task should be more independent from oral language proficiency than an oral reading fluency measure. Only a few studies examined the relation of silent reading fluency to comprehension in monolinguals, reporting moderate ( $r = 0.38$  in Fuchs et al., 2001) to strong correlations ( $r = 0.75$  in Klauda and Guthrie, 2008) for fourth and fifth graders, respectively. There appear to be no

studies on the relation of silent reading fluency to comprehension for bilinguals.

Finally, the nature of the fluency-comprehension relation varies developmentally. Reading fluency is expected to transition from word-based to word-integration fluency around age 11 (Berninger et al., 2010), including for L2 reading (Geva and Farnia, 2012). One study with fifth graders suggested that L2 reading fluency may contribute unique variance to reading comprehension beyond word reading fluency and oral language proficiency (Crosson and Lesaux, 2010), supporting a developmental shift to text fluency as a predictor of L2 reading comprehension. Geva and Farnia (2012) similarly found that by Grade 5, word and text fluency formed separate factors for both L1 and L2 readers, and that text fluency contributed uniquely to reading comprehension for both groups.

Around the same period reading shifts from oral to silent mode, which are independent forms of reading (Kim et al., 2011). The use of oral reading for fluency measures, then, may not be appropriate for readers at this age range when reading shifts to silent mode, and this may be particularly true for bilingual readers, as noted above. That is, bilingual readers' oral reading fluency may act simply as a proxy for oral language proficiency in general (e.g., Baker and Good, 1995), and therefore may underestimate children's written language ability. For adult English L2 learners, the correlation of oral and silent reading fluency varies with English proficiency, such that increased proficiency yields a closer correlation between the two reading modes (Lems, 2012).

The focus of the present study is to examine relations between reading fluency and comprehension in bilingual children across the age range where fluency shifts from word- to text-level. To circumvent the above-mentioned issues with oral reading fluency, we measured silent reading fluency of extended text passages in addition to comprehension and word fluency, decoding, and listening comprehension. To appraise silent reading fluency, we applied complexity measures of recurrence and fractal scaling to a self-paced reading task, using word reading times as a series. These complexity measures have been used as a means of quantifying aspects of the reading process, such as stability and structure (Wallot et al., 2012). They are thought to measure the degree to which text and language performance constrain reading (Wallot, 2015), where the better a reader can decode and comprehend a text, the text will have a more systematic influence on reading behavior, as seen in changes in reading process complexity. In previous studies with monolinguals, complexity measures from recurrence quantification and fractal analyses were shown to be a sensitive gauge of individual differences in silent reading fluency. O'Brien et al. (2014) reported that complexity measures of stability and orderliness for reading a text passage varied across age groups with increasing degrees of reading fluency. Wallot et al. (2014) further showed that the complexity measures were better predictors of reading comprehension than reading speed.

In the current study, we examine the relations of fluency and other factors to reading comprehension for bilingual children, across the ages where fluency is expected to shift from word to text level processing. We wanted to examine these relations while

keeping comprehension difficulty similar across ages, so we held passage difficulty constant by having children read grade-leveled passages. To investigate what contributes to reading proficiency and comprehension for our bilingual readers, we examined relations between decoding skill, listening comprehension, and fluency with reading comprehension following the simple view of reading. We might expect that decoding skills are relatively less important than listening comprehension within our bilingual sample. Further, we might expect that fluency measured for silent reading additionally contributes to reading comprehension beyond listening comprehension. However, the nature of fluency—comprehension relations may also vary by age, with text fluency becoming more relevant later on. Following these expectations, we address the following research questions:

1. What are the relative contributions of decoding and listening comprehension to reading comprehension within bilingual readers? Does text fluency act as a mediator to reading comprehension?
2. To what extent are text and word fluency related to reading comprehension for bilingual children?
3. Does the relative strength of the fluency-to-reading comprehension relation vary across age groups for word fluency compared with text fluency?

## METHODS

### Participants

One hundred and five children in grades 3, 4, and 5 participated ( $n$ 's = 33, 35 and 37, respectively) from two public schools in Singapore. The children varied in their home language backgrounds, including Mandarin, Malay, Tamil, and others, but all were schooled in English as well as their designated mother tongue. First and primary language use was ascertained with a self-report survey, which included items regarding frequency and domains of use per language, ranking of best language across modes, and family use. A majority of students reported English language (EL) as their best language overall (63), while almost  $\frac{1}{4}$  indicated Chinese (CL) was their best language (24) and 18 reported other languages as their best language [these included Malay (8), Tamil (4) and Burmese, Cantonese, Hokkien, Korean, Myanmar and Tagalog (6 altogether)].

### Measures

All assessments including the story reading task took about 1 h to complete and were given in one session at the child's school.

### Silent Reading Fluency For Text

Silent reading fluency for text was assessed with an experimental measure of story reading. Each individual read a grade-appropriate story in English that was rendered word-by-word on a MacBook Pro computer using a custom MatLab Psychophysics Toolbox script (Brainard, 1997). As the participant read the story, the words accumulated on the screen in a self-paced manner with a button press for each word (Just et al., 1982). After filling with text, the screen was refreshed for the next page of accumulating text. Response times for each word in the

passage were turned into a time-ordered series for submission to non-linear analyses. After reading the story, 10 multiple-choice comprehension questions were read aloud. Questions included literal, inferential, vocabulary from context, and main idea types. There was one story per grade level (based on "Clever Trevor" by Sarah Albee for P3, "Clever Beatrice" by Margaret Willey and Heather Solomon for P4, and "Fiona's Luck" by Teresa Bateman for P5). Stories were modified from published literature to be close to 1100 words long, as necessitated by the non-linear analyses. Details of the story lengths and readability indices are provided in **Table 1**.

Complexity measures, were derived using two different non-linear analytic methods: detrended fractal analysis (DFA), multifractal detrended fluctuation analysis (MFDFA), and recurrence quantification analysis (RQA). The appendix to this paper gives a more detailed overview over the three methods, including details of calculating each measure, prior applications, and current interpretations of—and hypotheses regarding—these measures in psychological research.

We used DFA to examine the fractal structure of the series of word reading response times across a text passage. DFA describes how variability changes across different time scales, with a fractal scaling exponent (i.e.,  $H$ , Hurst) that quantifies the degree of long-range correlation in the time series. This scaling exponent, which we refer to in this paper as *monofractal structure*, indicates whether word-by-word response times are independent of each other, or whether there are short-term correlations between response times (e.g., reading word <sub>$n$</sub>  affects the reading of adjacent word <sub>$n+1$</sub> ) or perhaps longer-term correlations across larger segments of the text (e.g., words within sentences or paragraphs or whole passages). From previous work (Kloos and Van Orden, 2010; Kuznetsov and Wallot, 2011; O'Brien et al., 2014; Wallot et al., 2014), we conceptualize that more proficient reading is guided and constrained by extraneous text features and therefore exhibits weaker long-term traces or links across word reading times during self-paced reading (O'Brien et al., 2014; Wallot et al., 2014) or fixation duration during reading (Wallot

**TABLE 1 | Characteristics of story texts.**

STORY	P3	P4	P5
Length (words)	1183	1105	1292
Average sentence length (words)	7.2	11.6	11.8
Average word length (letters)	3.9	3.9	4.1
<b>READABILITY INDEICES:</b>			
ATOS Grade level	2.4	4.4	5.3
Flesch-Kincaid Grade level	1.4	3.0	3.8
<b>WORD FREQUENCIES:</b>			
Graded Corpus (WFG)	1427.6	1537.7	1233.2
Total Corpus (WFG)	23062	26236	21108
Singapore Corpus (ICE)	427	382	480

ATOS (Accelerated Reader, 2011); Flesch-Kincaid (Coh-metrix, Graesser et al., 2004); Mean type frequencies according to WFG = Word Frequency Guide (Zeno et al., 1995); ICE = International Corpus of English: Singapore Corpus (Nelson, 2002). Stories correspond to the grade level to which they were administered (P3 for grade 3, P4 for grade 4, P5 for grade 5 students).

et al., 2015). Essentially, this means that more proficient, fluent reading is characterized by reduced scaling, with  $H$  relatively closer to random fluctuations or white noise.

A second variable we estimated is *multifractal structure*, using MF DFA, which is an expansion of DFA. Multifractal scaling captures the degree to which monofractal structure changes in the response-time series. Multifractal structure signifies that the series of reading times is heterogeneous and exhibits interactions across time-scales (Ihlen and Vereijken, 2010; Kelty-Stephen et al., 2013), for example where different levels of discourse (topical, syntactical, semantical, sub-lexical) interact with each other to guide readers' comprehension (Booth et al., 2016). It is expected that multifractal structure may capture a reader's adaptive behavior while reading a text for comprehension (Wallot et al., 2014). In this case, sudden on-line changes may occur during reading, perhaps reflecting a more dramatic change in understanding or insight (Stephen et al., 2009), rather than more gradual shifts where meaning is built cumulatively from preceding text (Donald, 2007).

The third non-linear method, RQA, quantifies recurrent patterns in the reading time series, describing the system's stability or orderliness. This analysis yields estimated parameters of the orderliness or recurrent patterning of a system within the task's phase space, quantified as the proportion of data points that are part of a recurring pattern, and referred to here as *%Determinism*. Prior studies show that reading performance becomes more structured with higher determinism as reading skill increases (Wijnants et al., 2009, 2012; Wallot et al., 2012), and determinism is higher in more fluent readers (O'Brien et al., 2014; Wallot et al., 2014). It has been suggested that measures of temporal structure of reading times, such as RQA *%Determinism*, capture how well readers utilize the informational structure of a text, and conversely, how well a text constrains the reading process toward efficient and fluent reading (Wallot, 2015, 2016). Hence, high degrees of *%Determinism* should be positively correlated with aspects of reader skill. In that sense, *monofractal structure* and *%Determinism* are conceptually closely related, but there are differences as well (see Appendix).

Thus, for our analysis we include three complexity measures to capture aspects of structure and orderliness of reading times over the series of the text: monofractal structure, multifractal structure, and *%Determinism*.

## Word Reading Fluency

Word reading fluency was assessed with the TOWRE (Test of Word Reading Efficiency) sight word subtest (Torgesen et al., 1999). This test involves sight word reading of a list of words, starting from more to less frequent. Scores are tallied as number of correctly read words within the 45 s time limit.

## Decoding

Decoding was assessed with the TOWRE phonemic decoding subtests (Torgesen et al., 1999). Similar to the above subtest, this test involves reading of a list of non-words by phonemically decoding them, and scores are tallied as the number of correctly read non-words within the 45 s time limit.

## Reading Component Skills

Reading component skills including phonological awareness and rapid symbol naming, two robust predictors of reading ability, were also assessed and were used descriptively. *Phonological awareness* was measured with the CTOPP (Comprehensive Test of Phonological Processing) Elision subtest (Wagner et al., 1999) which is a phoneme deletion task. *Rapid symbol naming* was assessed with the RAN/RAS (Rapid Automatized Naming and Rapid Alternating Stimulus Tests) letters subtest (Wolf and Denckla, 2004).

## Oral Language Proficiency

Oral language proficiency was assessed with the WJIII (Woodcock-Johnson III) Listening Comprehension Cluster, which is comprised of the WJIII-Understanding Directions and Oral Comprehension subtests (Woodcock et al., 2007). Understanding directions involves listening to increasingly complex sequences of instructions and responding by pointing to objects in a picture. Oral comprehension involves listening to short passages and using semantic/syntactic cues to supply missing words within the passage.

## Data Preparation

Prior to FA, extreme response times of 10 s or longer were removed, a threshold adapted to children's reading times (O'Brien et al., 2014). On average, 3.7 data points were eliminated per participant ( $SD = 4.0$ ), which amounted to 0.31% of all data points. The extreme scores were removed because they can distort the fractal analysis, while the slight disruption on the series' time order has minimal impact, as long as a minimum of 1024 observations are maintained (Holden, 2005, p. 285–287). There are several methods available to estimate the scaling relations, including spectral analysis (SA), standardized dispersion analysis (SDA), and detrended fluctuation analysis (DFA). DFA results are reported here (Peng et al., 1995), and were corroborated with the two other methods of FA. To assess multifractal structure, multifractal detrended fluctuation analysis (MF DFA) was used (Kantelhardt et al., 2002).

For RQA, all data in the time-series were entered into the analysis using the Commandline Recurrence Plots software (Marwan, 2011). Data are first rescaled relative to the Euclidian distance separating points in reconstructed phase space (using time-delayed copies of the time series as surrogate dimensions) providing an intrinsically scaled metric across the set of data. *%Determinism* was calculated using parameters of embedding delay (1), dimension (5) and radius (0.4) following procedures described in Webber and Zbilut (2005).

For the standardized assessments of word reading, decoding, phonological awareness, rapid naming, and listening comprehension Student's  $t$ -statistic was used for the correlation and multiple regression analyses given the difference in the current sample from the normative sample.

## RESULTS

### Descriptives of Reading Skills

Average performance per grade level on the standard reading and language measures are shown in **Table 2**, along with the



**TABLE 2 | Mean (SE) standard scores on reading and language tests.**

Test	Grade 3 (n = 33)	Grade 4 (n = 34)	Grade 5 (n = 37)	Overall (N = 105)
<b>READING COMPONENT SKILLS</b>				
Rapid Naming	111.4 (2.6)	110.0 (2.8)	118.7 (2.3)	113.5 (1.5)
Phonological awareness	8.3 (0.3)	8.0 (0.5)	8.9 (0.5)	8.4 (0.2)
Decoding	104.8 (2.4)	107.1 (2.4)	111.2 (1.6)	107.8 (1.2)
Word reading fluency	111.9 (2.1)	110.0 (1.9)	93.6 (2.1)	111.7 (1.1)
Listening Comprehension	81.9 (2.7)	82.9 (2.6)	89.3 (2.4)	84.8 (1.5)

Rapid Naming (RAN/RAS), Decoding (TOWRE-phonemic decoding), Word reading fluency (TOWRE-sight words), and Listening comprehension (WJIII) standardized scores are relative to a mean of 100 (SD = 15). Phonological awareness (CTOPP-Elision) is a scale score with a mean of 10 (SD = 3).

overall sample mean. Standard scores are presented here, to give an impression of peer-referenced skill levels across grades, but it should be noted that the standardized scores are based on published normative data from monolingual English speakers. As can be seen most of the averages are within the normal range, with the exception of Listening Comprehension, which is about 1 SD or more below the mean for P3 and P4 groups.

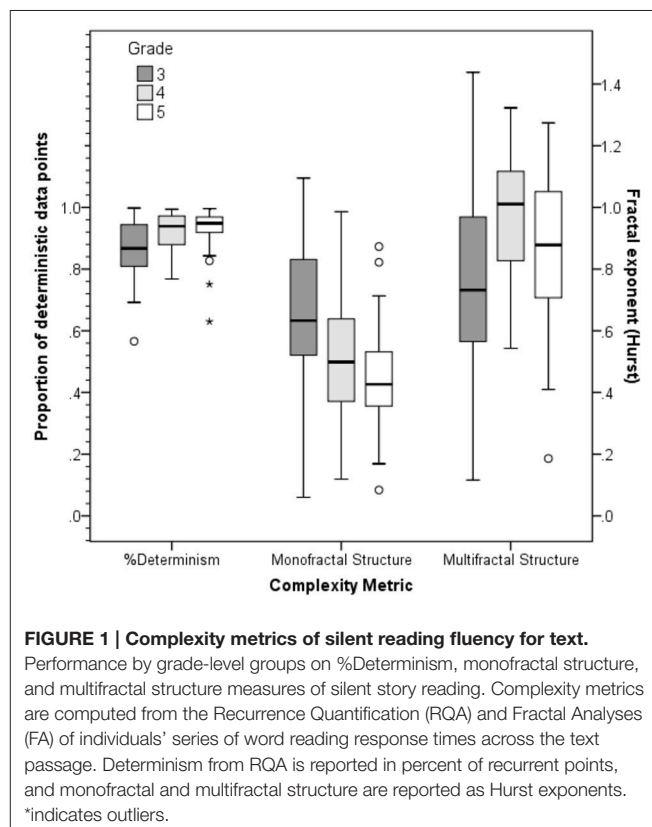
Performance on the silent passage reading task is presented in **Table 3** for reading rate (wpm) and comprehension scores and the complexity measures. Notably text comprehension did not differ across the three grade level groups [ $F_{(2, 101)} = 1.70$ ,  $p = 0.189$ ,  $\eta^2 = 0.032$ ], most likely as a result of matching the text difficulty appropriately for each grade. Reading speed, on the other hand, did increase significantly across the grades [ $F_{(2, 101)} = 9.72$ ,  $p < 0.001$ ,  $\eta^2 = 0.161$ ], with 3rd graders reading slower than 4th and 5th graders ( $p$ 's  $< 0.001$ ), but no difference for 4th and 5th graders in reading speed ( $p = 0.783$ ) according to Bonferroni corrected *post-hoc* tests.

For the complexity measures applied to the text reading times series, there was a significant increase in %Determinism of reading times across grades [ $F_{(2, 101)} = 7.63$ ,  $p < 0.001$ ,  $\eta^2 = 0.131$ ], indicating that word reading times became more regular for older readers. Bonferroni corrected *post-hoc* tests revealed that 3rd graders showed lower %Determinism of reading times than 4th ( $p = 0.003$ ) and 5th graders ( $p = 0.002$ ), but 4th and 5th graders did not differ in %Determinism ( $p = 0.929$ ). Monofractal structure in reading times also differed across grades [ $F_{(2, 101)} = 10.21$ ,  $p < 0.001$ ,  $\eta^2 = 0.168$ ]. *Post-hoc* tests with Bonferroni correction showed that the 3rd grade group had a greater fractal exponent than both 4th ( $p = 0.003$ ) and 5th graders ( $p < 0.001$ ). This differs from the previous finding with monolingual children, who showed no age effects of monofractal structure (O'Brien et al., 2014). Furthermore, we analyzed the change of multifractal structure in reading times, which increased with grade  $F_{(2, 101)} = 5.65$ ,  $p = 0.005$ ,  $\eta^2 = 0.101$ . Bonferroni corrected *post-hoc* tests revealed that 3rd graders showed less multifractal structure in reading times compared to 4th graders ( $p = 0.004$ ). No other effects were apparent (both  $p > 0.133$ ). **Figure 1** shows the group means for each of the three complexity metrics.

**TABLE 3 | Performance measures on the silent passage reading task.**

Test	Grade 3 (n = 33)	Grade 4 (n = 34)	Grade 5 (n = 37)	Overall (N = 105)
Reading Rate (WMP)	100.2 (32.6)	132.3 (50.4)	134.8 (42.1)	123.0 (45.0)
Story Comprehension	68% (20–100)	76% (20–100)	77% (30–100)	74% (20–100)
<b>COMPLEXITY MEASURES</b>				
%Determinism	0.86	0.92	0.93	0.90
Monofractal Structure	0.66	0.49	0.44	0.52
Multifractal Structure	0.76	0.97	0.85	0.86

Silent reading rate is reported in words per minute, and Story comprehension is percent correct, with the group's range shown below.

**FIGURE 1 | Complexity metrics of silent reading fluency for text.**

Performance by grade-level groups on %Determinism, monofractal structure, and multifractal structure measures of silent story reading. Complexity metrics are computed from the Recurrence Quantification (RQA) and Fractal Analyses (FA) of individuals' series of word reading response times across the text passage. Determinism from RQA is reported in percent of recurrent points, and monofractal and multifractal structure are reported as Hurst exponents. \*Indicates outliers.

To examine relations across all measures, including traditional literacy and language proficiency tests as well as the silent reading task, we calculated zero-order Pearson-correlations with age partialled out. For the standardized tests (CTOPP, TOWRE, WJIII) we used Student's *t*-statistic based on the sample's mean and standard deviation. From **Table 4**, it appears that rapid naming had a stronger relation to decoding and word fluency skills than phonological awareness, and a small correlation with monofractal structure. Phonological awareness was not systematically related to the complexity measures. Decoding and word fluency skills, on the other hand, were related to each other, and showed similar relations with monofractal structure and listening comprehension, but only word fluency was correlated

**TABLE 4 | Correlations between reading and language measures for the whole sample.**

	2.	3.	4.	5.	6.	7.	8.	9.
1. Rapid Naming	0.019	−0.368	−0.675	−0.166	0.304*	−0.073	−0.136	−0.169
2. Phonological awareness	1	0.247*	0.043	−0.013	−0.076	−0.158	0.101	−0.014
3. Decoding		1	0.577	0.220*	−0.384	−0.056	0.274	0.154
4. Word reading fluency			1	0.231*	−0.417	0.085	0.296	0.259*
5. %Determinism				1	−0.345	0.472	0.271	0.037
6. Monofractal structure					1	0.128	−0.326	−0.310
7. Multifractal structure						1	0.025	0.004
8. Listening comprehension							1	0.534
9. Reading Comprehension								1

Age Partialled Out. \*marks  $p < 0.05$ ; bold marks  $p < 0.01$ . Measures include Student's  $t$ -statistic for rapid naming (RAN letters subtest), phonological awareness (CTOPP Elision subtest), decoding efficiency (TOWRE phonemic decoding subtest), and word reading fluency (TOWRE sight word subtest), and listening comprehension (WJIII), and percent correct for reading comprehension. Complexity measures for silent reading fluency include %Determinism, and mono- and multifractal scaling exponents.

with reading comprehension. Further, while the complexity measures for silent text reading showed some interrelationships, it is only monofractal structure that showed a relation to reading comprehension scores. This correlation was stronger than that between word fluency and reading comprehension. Listening comprehension showed strongest correlation with reading comprehension for this sample.

## Relation of Components of the Simple View of Reading

To address the first research question hierarchical regression models were run with reading comprehension as the criterion measure. Age was entered as the first step, then decoding was entered into the second step and listening comprehension scores into the final step. Overall the model accounted for 29% variance in reading comprehension (see **Table 5**). Decoding did not contribute significantly, but listening comprehension did, accounting for 25% unique variance in reading comprehension after accounting for the other variables. When the order of decoding and listening comprehension predictors was reversed, listening comprehension still accounted for significant variance and decoding did not contribute any additional variance. This confirms the first hypothesis that listening comprehension would be a more potent factor for reading comprehension compared with decoding in our bilingual sample.

To test the second prediction that fluency plays a mediating role for reading comprehension and either decoding or listening comprehension, mediation analysis was run using structural equation modeling (Lavaan statistics package within R, Rosseel, 2012). First, a model of reading comprehension scores with decoding as the predictor and one of the text fluency measures as mediator was run. Only the model with monofractal structure entered as mediator showed a significant indirect effect (indirect effect  $Z = 2.49$ ,  $p = 0.013$ , direct effect  $Z = 0.33$ ,  $p = 0.74$ ,  $R^2 = 0.115$ ). The models with %Determinism and multifractal structure showed no significant effects, either direct or indirect, of decoding on reading comprehension ( $R^2$ 's = 0.02). Second, models of reading comprehension regressed on listening comprehension revealed significant direct effects with

%Determinism or multifractal structure as a mediator ( $Z = 6.3$ ,  $p$ 's  $< 0.01$ ,  $R^2$ 's = 0.27). Only monofractal structure showed a trend toward a significant mediation effect for the listening and reading comprehension relation (indirect effect  $Z = 1.89$ ,  $p = 0.059$ , direct effect  $Z = 5.45$ ,  $p < 0.01$ ,  $R^2 = 0.311$ ).

Thus, in this current bilingual sample, the measure of decoding skill was a much weaker predictor of reading comprehension than the measure of listening comprehension skill, and only showed an indirect effect on reading comprehension through text fluency (monofractal structure). Listening comprehension, on the other hand, was directly related to reading comprehension, and only monofractal structure for text fluency showed a tendency to mediate this relation.

## Relation of Word Level and Text Level Fluency to Reading Comprehension

To address the second research question hierarchical regression models were run with reading comprehension as the criterion measure and fluency measures as predictors. Age was entered as the first step, then word reading fluency was entered into the second step. The three complexity measures (%Determinism, Monofractal structure, and Multifractal structure) were entered into the last step of the model. Overall the model accounted for almost 14% variance in reading comprehension (see **Table 6**). The model with word fluency and age tended toward significance ( $p = 0.06$ ), whereas the addition of the text fluency variables showed a significant change in explained variance for the model. Of the three text fluency measures, monofractal structure accounted for 8% unique variance, while contribution of variance from the other complexity measures was not significant. For the reverse order of entry, with text fluency measures entered in the second step and word fluency in the final step, text fluency, and age accounted for 11% of the variance in comprehension, and word fluency did not add significant variance beyond this.

## Word and Text Fluency across Grade

Next we addressed the third research question, and the prediction that fluency shifts from word level to text level around

**TABLE 5 | Hierarchical Multiple regression predicting reading comprehension from decoding and listening comprehension.**

Step	Variables entered	$\Delta R^2$	$\beta$	$p$	Variables entered	$\Delta R^2$	$\beta$	$p$
1	Age	0.021	0.146	0.14	Age	0.012	0.146	0.14
2	Age	0.023	0.153	0.12	Age	<b>0.287</b>	0.164	0.05
	Decoding		0.153	0.12	Listening comprehension		0.529	0.00*
3	Age	<b>0.256</b>	0.165	0.05	Age	<b>0.280</b>	0.165	0.05
	Decoding		0.008	0.93	Listening comprehension		0.527	0.00*
	Listening comprehension		0.527	0.00*	Decoding		0.008	0.93

Bold indicates significant change in explained variance. Listening comprehension (WJIII) and Decoding (TOWRE), predictors entered as Student's  $t$ -statistic. \*indicates significant effect of predictor.

**TABLE 6 | Hierarchical Multiple regression predicting reading comprehension from reading fluency measures.**

Step	Variables entered	$\Delta R^2$	$\beta$	$p$	Variables entered	$\Delta R^2$	$\beta$	$p$
1	Age	0.021	0.146	0.14	Age	0.021	0.146	0.14
2	Age	0.035	0.160	0.10	Age	<b>0.110</b>	0.034	0.74
	Word fluency		0.186	0.06	%Determinism		−0.156	0.21
					Monofractal		−0.397	0.00*
					Multifractal		0.125	0.28
3	Age	<b>0.080</b>	0.053	0.62	Age	0.005	0.053	0.62
	Word fluency		0.077	0.47	%Determinism		−0.167	0.19
					Monofractal		−0.365	0.00*
					Multifractal		0.126	0.28
	%Determinism		−0.167	0.19	Word fluency		0.077	0.47
	Monofractal		−0.365	0.00				
	Multifractal		0.126	0.28				

Bold indicates significant change in explained variance. Listening comprehension (WJIII) and Decoding (TOWRE), predictors entered as Student's  $t$ -statistic. \*indicates significant effect of predictor.

**TABLE 7 | Relation between reading fluency measures and reading comprehension for Grade 3.**

	2.	3.	4.	5.
1. Word Fluency	0.415*	<b>−0.621</b>	−0.063	<b>0.535</b>
2. %Determinism	1	−0.406*	0.431*	−0.128
3. Monofractal		1	0.259	<b>−0.510</b>
4. Multifractal			1	−0.299
5. Reading Comprehension				1

\*marks  $p < 0.05$ ; bold marks  $p < 0.01$ .

**TABLE 8 | Relation between reading fluency measures and reading comprehension for Grade 4.**

	2.	3.	4.	5.
1. Word Fluency	0.367*	−0.346*	0.190	0.019
2. %Determinism	1	−0.296	0.279	−0.095
3. Monofractal		1	0.072	−0.098
4. Multifractal			1	−0.038
5. Reading Comprehension				1

\*marks  $p < 0.05$ .

fourth grade. It was of interest to see whether there were different patterns amongst these measures at any juncture across the hypothesized developmental shift from word- to text-reading. Intercorrelations for each grade are shown separately in **Tables 7–9**. As predicted, the influence of word reading fluency on reading comprehension declined with age, and was only significantly correlated in the third grade group. The measures of text fluency, on the other hand, showed different patterns of variation over the three age groups. Monofractal structure, like

word fluency, was correlated with reading comprehension only for the P3 group, showing no significant relation for fourth and fifth grade children. Multifractal structure and %Determinism showed the opposite pattern, whereby they were not significantly related to reading comprehension in grade 3 or 4 groups, but were related in the fifth grade children.

Finally, we regressed reading comprehension on the fluency measures using stepwise multiple regression with a forward selection procedure. This allowed us to examine which of the

**TABLE 9 | Relation between reading fluency measures and reading comprehension for Grade 5.**

	2.	3.	4.	5.
1. Word Fluency	−0.102	−0.246	0.154	0.189
2. %Determinism	1	−0.109	<b>0.602</b>	0.335*
3. Monofractal		1	0.148	−0.186
4. Multifractal			1	0.398*
5. Reading Comprehension				1

\*marks  $p < 0.05$ ; bold marks  $p < 0.01$ .

fluency measures improved the prediction model best within each grade, although results are viewed with caution because the sample size per grade was small. The regressions confirmed the above observations from the correlation tables. For grade 3, significant predictors of reading comprehension included both word fluency (semipartial correlations,  $r = 0.373$ ) and the text fluency measures of monofractal structure ( $r = -0.313$ ) as well as %Determinism ( $r = -0.442$ ) ( $R^2 = 0.485$ ,  $F = 6.1$ ,  $p = 0.02$ ). For grade 4, none of the fluency measures contributed significantly to the prediction of reading comprehension, whereas for grade 5 only multifractal structure of text fluency was a significant predictor ( $r = 0.398$ ,  $R^2 = 0.134$ ,  $F = 6.4$ ,  $p = 0.016$ ). Thus, there was an overall shift from word to text fluency over these grade levels.

## DISCUSSION

The current study examined the relation between silent reading fluency and comprehension for bilingual children across the age range where fluency is proposed to shift from word-level to text-level. While some of the findings replicate those with monolingual readers, there were also some differences with regard to the interrelations of skills and processes with reading comprehension, and to age-related variations in fluency of silent text reading.

Interrelations between the fluency measures and basic reading related skills showed that rapid naming, but not phonological awareness, was related to both word fluency and silent text reading fluency. This follows from prior monolingual research where rapid naming is a better predictor of word fluency, and phonological awareness is better for predicting word reading accuracy (e.g., Schatschneider et al., 2004). Interestingly, rapid naming was related to the text fluency metric of monofractal structure, which, unlike word fluency, differs from RAN in that it is not a simple rate measure, but an indicator of the structure of reading times across the text.

The text fluency metrics of %Determinism and monofractal structure were also related to skills of decoding and word fluency, and all four of these measures were correlated with listening comprehension. Decoding and listening comprehension tend to show stronger correlations within monolingual samples ( $r$ 's = 0.40 to 0.60, Foorman et al., 2015) compared with here ( $r = 0.27$ ), implying that these skills may be more dissociated or develop more independently in bilingual readers (e.g., Jackson and Lu, 1992).

Reading comprehension, on the other hand, was not significantly related to decoding, in contrast to findings with monolingual readers of similar age (Foorman et al., 2015). Only listening comprehension, along with word fluency and the text fluency measure of monofractal structure, were significantly correlated with reading comprehension. Listening comprehension was a significant predictor of reading comprehension, explaining 25% unique variance beyond age and decoding skills, and showing a direct effect on reading comprehension. Decoding skills did not show such an impact, but was only related to reading comprehension indirectly, mediated by text fluency (monofractal structure). This further supports the prediction that decoding would play a lesser role in reading comprehension for bilingual readers. The findings also support the role of fluency as a mediator between decoding and comprehension, similar to findings with monolingual readers (Silverman et al., 2013), and suggests that issues with poor comprehension are more related to fluency than decoding for bilingual readers (e.g., Crosson and Lesaux, 2010; Chen et al., 2012; Geva and Farnia, 2012).

Language skills are found to relate to text fluency for monolingual readers (Cutting et al., 2009), and even moreso for second language learners (Geva and Zadeh, 2006; Crosson and Lesaux, 2010). Moreover, individual differences in language skills contribute more to text fluency than word level fluency does, and particularly for second-language or bilingual readers (Geva et al., 1997; Buly and Valencia, 2002; Geva and Farnia, 2012). While most of the children in this study rated English as their best language (probably a result of English being the main language of instruction throughout primary school), about 40% reported their mother tongue as their better language. Within this mixed group of bilinguals, neither their age of acquisition of English (by or after 3 years of age) nor their first language status (English or other language learned first) had any bearing on their text fluency performance, as indicated by between groups comparisons. But regardless of these factors, oral proficiency in English was related to their reading fluency performance, according to the correlational and regression analyses.

With regard to the contribution of text fluency to reading comprehension, it was shown that monofractal structure contributed significant variance after controlling for age, whereas multifractal structure and %Determinism did not. This contrasts with Wallot et al. (2014) wherein %Determinism was the best predictor of comprehension, while monofractal structure added unique variance only for oral and not for silent reading. Methodological differences in that study, including use of a single story matched to the youngest readers (grade 2), may explain the difference in findings. That is, while participants in Wallot et al. (2014) received the same texts and accordingly showed increases in comprehension score with age, participants in the present study received text of age-matched difficulty and accordingly did not show changes in comprehension scores with age. Alternatively, it may indicate a difference in the manner by which skilled, fluent reading emerges in monolingual vs. bilingual readers. That is, the way the reading system is assembled to perform the task of comprehending text may differ between monolingual and bilingual readers. For bilingual readers overall,

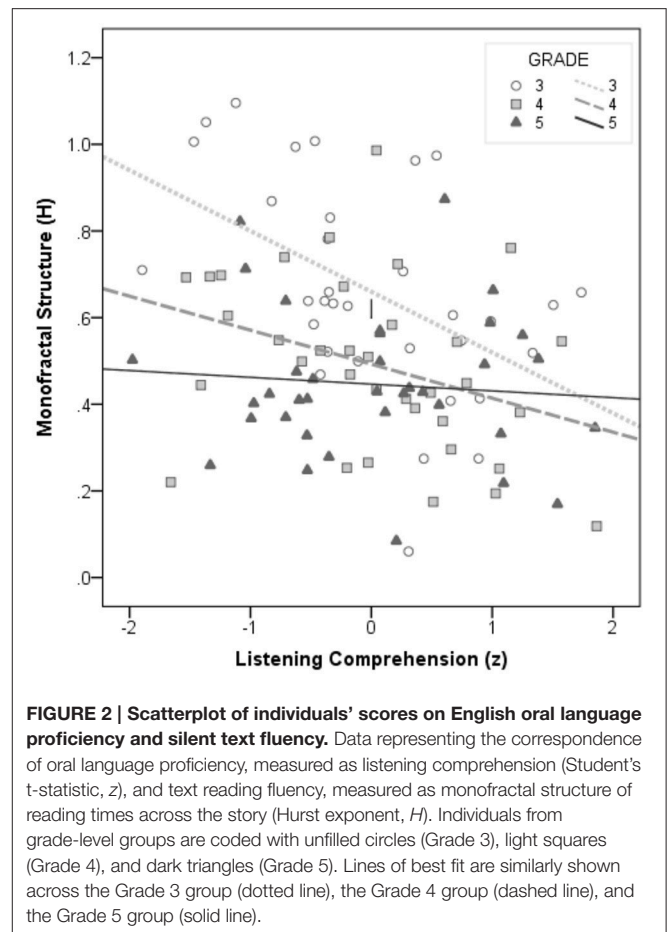


the optimal state for comprehension may be a tight coupling of reading-time performance to the ongoing informational input provided by the text. Further study is required to confirm these ideas, especially given the age related differences we found.

Nonetheless, the finding that processes for silent reading fluency, as indicated by monofractal structure here, contributes to reading comprehension coincides with earlier findings with English language learners. Oral text fluency was found to uniquely predict reading comprehension by Crosson and Lesaux (2010) with grade 5 ELLs after controlling for word reading fluency and oral language proficiency, and by Jeon (2012) with adult ELLs after controlling for word reading fluency and pseudoword reading. The present findings with bilingual children show that monofractal structure for text fluency and listening comprehension for oral language proficiency are the strongest predictors of reading comprehension of the stories. These two factors are also strongly related to each other. Their negative relation, as seen in the scatterplot (Figure 2) shows that those with stronger language proficiency also show decreased monofractal structure of their reading times (indicating increased text fluency). When broken down by age groups, we can see that this strong relation may be driven primarily by the youngest group of third graders. By grade 5, it appears language proficiency has no bearing on the text fluency metric. Further examination of the relation of both factors to reading comprehension did not support a model where text fluency mediates the relation between listening comprehension and reading comprehension (e.g., Kim and Wagner, 2015). Instead, both text fluency and listening comprehension contributed unique variance to reading comprehension when the other variable was controlled.

With regard to word level compared with passage level fluency, text fluency showed a stronger relation to reading comprehension than word fluency, though word fluency did show some correlation with comprehension despite being measured with a separate task. Of the text fluency measures, only monofractal structure showed a significant correlation and predicted unique variance in reading comprehension (8%). This is similar to the results with monolingual readers reported in Wallot et al. (2014). The relation of the monofractal structure measure and reading comprehension is negative, as found earlier, suggesting that better reading is a consequence of processes that are more strongly driven by the structure of the text. However, the corollary finding, that better reading is also characterized by cognitive reorganization during reading as indexed by a positive relation of the multifractal structure with reading comprehension, was not replicated in this sample. There was variation across the age groups, however, with regard to relations between the fluency metrics and reading comprehension. Differences across age are informative, as these were not examined previously because the sample size was smaller (Wallot et al., 2014).

For the complexity measures of silent text reading fluency, the bilingual readers showed the same pattern in %Determinism across grade level as monolingual readers of English (O'Brien et al., 2014). This measure of the degree of order in reading times across the story text showed increasing structure in



**FIGURE 2 | Scatterplot of individuals' scores on English oral language proficiency and silent text fluency.** Data representing the correspondence of oral language proficiency, measured as listening comprehension (Student's *t*-statistic, *z*), and text reading fluency, measured as monofractal structure of reading times across the story (Hurst exponent, *H*). Individuals from grade-level groups are coded with unfilled circles (Grade 3), light squares (Grade 4), and dark triangles (Grade 5). Lines of best fit are similarly shown across the Grade 3 group (dotted line), the Grade 4 group (dashed line), and the Grade 5 group (solid line).

reading performance with age. In the previous study, second graders showed less determinism compared with fourth and sixth graders, who in turn showed less determinism in reading times than adult readers. In the present study, the shift to greater %Determinism similarly occurred at grade 4.

We also found a difference across age groups in monofractal structure of children's reading times. Third grade children had greater monofractal exponents (*H*) compared with the older children. This differs from previous results (O'Brien et al., 2014), wherein the monolingual readers showed no age related differences in monofractal structure. In that study, one story rated as grade 2.5 (ATOS) was given to all age groups, effectively inducing greater levels of fluent reading with increasing age. Here, the reading texts were roughly matched to age groups: the 3rd grade group read a 2.4 ATOS level story, while the other groups read 4.4 and 5.3 ATOS leveled texts. While word lengths were similar across the three stories, sentences were on average longer for the 4th and 5th graders' stories. Perhaps the more complex language represented in the higher level texts allowed for, or demanded more, attention to text structure, yielding lower fractal scaling that is more constrained by faster time scales with less long-range dependencies, closer to random fluctuations that are expected to reflect a closer constraining by the text.

Age had a differential effect on multifractal scaling than it did on monofractal scaling—showing an increase in the former and a decrease in the latter across the grade level groups. This finding follows the predicted outcome that more fluent reading is related to a constraining of performance to faster time scales driven by text structure, as reflected in smaller scaling exponents in monofractal structure, whereas it is also characterized by a higher degree of adaptive changes as the reader processes the meaning of the text, as reflected in larger exponents in multifractal structure. A larger multifractal exponent indicates that, whereas reading is constrained by small timescale features of the text (e.g., word by word, or within-phrase features), the reader is also attuned to larger timescale features (e.g., in the plot or setting of the story).

Examining the correlational analyses separately for each grade, we found that the relation of comprehension to monofractal scaling was significantly negative only at grade 3, whereas the relation to multifractal scaling was positive and only significant at grade 5. %Determinism showed the same pattern as multifractal scaling, where the relation to comprehension was only significant by grade 5. The developmental differences show that these metrics for “silent fluency” may capture different aspects of what we mean by fluency—as primarily text-driven speed earlier on, but with a later emphasis on order and also adaptive aspects of fluency that contribute to better comprehension.

It should be noted that reading comprehension did not differ across the age groups, so these changes in relation to fluency aspects are not simply due to improved comprehension generally. Further, although the older readers showed both decreased monofractal and increased multifractal structure and determinism as a group, it appears that individual differences in reading comprehension were only related to the multifractal structure and determinism for the P5 group. For the P3 children, the better comprehenders looked more like the older groups, with lower monofractal structure than their peers. That is, for the older readers good comprehension may act as a dynamic attractor state (as indicated by higher %Determinism) where current processing is constrained by what has already been read, but is also responsive to how well new information is integrated with previous context (as indicated by higher multifractal structure) (e.g., Paulson, 2005). For the younger readers, on the other hand, good comprehension appears to coincide with reading activity focused at small timescales (e.g., single word level, as indicated by lower monofractal structure), and follows the concept that these readers are still at the stage where they are “glued to the print” (Chall, 1996). At this stage, the difference between better and

poorer comprehenders is leveled at processing at the small time scale (e.g., word level recognition or decoding), and is not yet dependent on the readers' attunement to larger timescale features (e.g., meaning-based processing of the story). This is supported by the finding that word fluency showed a significant relation to reading comprehension only for the grade 3 group. By grade 4 and 5 it appears word fluency is no longer as important for comprehension, as text fluency becomes more prominent by grade 5 for our bilingual sample. Geva and Farnia (2012) similarly found that word level fluency only contributed to text fluency in early primary school, but by grade 5 text fluency became more aligned with language skills for both first language and second language learners.

In sum, for the bilingual readers we observed across the middle primary grades, the present results indicate that text fluency measured for silent reading predicted story reading comprehension, and that English language proficiency was also predictive of both reading fluency and reading comprehension performance. The present set of results should be treated with caution, as the sample size was small for examining predictive relations within each grade level. Further, our skills measures are based on single measurements rather than latent variables, and findings may be particular to the specific assessments we used. More research on the roles of fluency and oral language in bilingual reading is warranted, particularly given the apparent age-related variations in the relation between these skills.

## AUTHOR CONTRIBUTIONS

BO and SW designed the study. BO collected and analyzed the data. BO and SW interpreted the data and wrote the manuscript.

## ACKNOWLEDGMENTS

Preparation of this article was supported by the NIE Office of Education Research grant # SUG2812OBA. The authors wish to thank the school principals and heads of department, as well as the students who participated in this research. A portion of the study was presented at the 2013 European Conference on Developmental Psychology in Lausanne, Switzerland.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.01265>

## REFERENCES

- Accelerated Reader (2011). *Book Finder US*. Available online at: <http://www.arbookfind.com/bookdetail.aspx?q=59508&l=EN&slid=194455065> (Accessed August 15, 2011).
- Adolf, S., Catts, H., and Little, T. (2006). Should the simple view of reading include a fluency component? *Read. Writ.* 19, 933–958. doi: 10.1007/s11145-006-9024-z
- Baker, S. K., and Good, R. (1995). Curriculum-based measurement of English reading with bilingual Hispanic students: a validation study with second grade students. *School Psychol. Rev.* 24, 561–578.
- Bashir, A. S., and Hook, P. E. (2009). Fluency: a key link between word identification and comprehension. *Lang. Speech Hear. Serv. Schools* 40, 196–200. doi: 10.1044/0161-1461(2008/08-0074)
- Berninger, V. W., Abbott, R. D., Trivedi, P., Olson, E., Gould, L., Hiramatsu, S., et al. (2010). Applying the multiple dimensions of reading fluency to assessment and instruction. *J. Psychoeduc. Assess.* 28, 3–18. doi: 10.1177/0734282909336083
- Bialystok, E., Luk, G., Peets, K. F., and Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism* 13, 525–531. doi: 10.1017/S1366728909990423

- Booth, C., Brown, H. L., Eason, E. G., Wallot, S., and Kelty-Stephen, D. G. (2016). Expectations on hierarchical scales of discourse: multifractality predicts both short- and long-range effects of violating gender expectations in text-reading. *Discourse Process*. doi: 10.1080/0163853X.2016.1197811
- Brainard, D. H. (1997). The Psychophysics toolbox. *Spatial Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Buly, M. R., and Valencia, S. W. (2002). Below the bar: profiles of students who fail state Reading assessments. *Edu. Eval. Policy Anal.* 24, 219–239. doi: 10.3102/01623737024003219
- Caccia, D. C., Percival, D., Cannon, M. J., Raymond, G., and Basingthwaite, J. B. (1997). Analyzing exact fractal time series: evaluating dispersal analysis and rescaled range methods. *Phys. A* 246, 609–632. doi: 10.1016/S0378-4371(97)00363-4
- Chall, J. S. (1996). *Stages of Reading Development (2nd Edn.)*. Fort Worth, TX: Harcourt-Brace.
- Chen, X., Geva, E., and Schwartz, M. (2012). Understanding literacy development of language minority students: an integrative approach. *Read. Writ.* 25, 1797–1804. doi: 10.1007/s11145-012-9400-9
- Ciuciu, P., Abry, P., and He, B. J. (2014). Interplay between functional connectivity and scale-free dynamics in intrinsic fMRI networks. *Neuroimage* 95, 248–263. doi: 10.1016/j.neuroimage.2014.03.047
- Costa, A., Roelstraete, B., and Hartsuiker, R. J. (2006). The lexical bias effect in bilingual speech production: evidence for feedback between lexical and sublexical levels across languages. *Psychon. B. Rev.* 13, 972–977. doi: 10.3758/BF03213911
- Crosson, A. C., and Lesaux, N. K. (2010). Revisiting assumptions about the relationship of fluent reading to comprehension: Spanish-speakers text-reading fluency in English. *Read. Writ.* 23, 475–494. doi: 10.1007/s11145-009-9168-8
- Cutting, L. E., Materek, A., Cole, C. A. S., Levine, T. M., and Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Ann. Dyslexia* 59, 34–54. doi: 10.1007/s11881-009-0022-0
- DeRamirez, R. D., and Shapiro, E. S. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms. *School Psychol. Rev.* 35, 356–369.
- Donald, M. (2007). The slow process: a hypothetical cognitive adaption for distributed cognitive networks. *J. Physiol. Paris* 101, 214–222. doi: 10.1016/j.jphysparis.2007.11.006
- Foorman, B. R., Koon, S., Petscher, Y., Mitchell, A., and Truckenmiller, A. (2015). Examining general and specific factors in the dimensionality of oral language and reading in 4<sup>th</sup>–10<sup>th</sup> grades. *J. Educ. Psychol.* 107, 884–899. doi: 10.1037/edu0000026
- Fuchs, L. S., Fuchs, D., Hosp, M. K., and Jenkins, J. R. (2001). Text fluency as an indicator of reading competence: a theoretical, empirical, and historical analysis. *Sci. Stud. Read.* 5, 239–256. doi: 10.1207/S1532799XSSR0503\_3
- Geva, E., and Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL. *Read. Writ.* 25, 1819–1845. doi: 10.1007/s11145-011-9333-8
- 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
- Geva, E., and Zadeh, Y. Z. (2006). Reading efficiency in native English-speaking and English-as-a-second-language children: the role of oral proficiency and underlying cognitive-linguistic processes. *Sci. Stud. Read.* 10, 31–58. doi: 10.1207/s1532799xssr1001\_3
- Geva, E. (2006). “Second-language oral proficiency and second-language literacy,” in *Developing Literacy in Second-language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth*, eds D. August and T. Shanahan (New York, NY: Routledge), 123–140.
- Goldberger, A. L., Amaral, L. A., Hausdorff, J. M., Ivanov, P. C., Peng, C. K., and Stanley, H. E. (2002). Fractal dynamics in physiology: alterations with disease and aging. *Proc. Natl. Acad. Sci. U.S.A.* 99, 2466–2472. doi: 10.1073/pnas.012579499
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., and Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Mem. Cogn.* 33, 1220–1234. doi: 10.3758/BF03193224
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., and Cai, Z. (2004). Coh-metrix: analysis of text on cohesion and language. *Behav. Res. Methods Instrum. Comput.* 36, 193–202. doi: 10.3758/BF03195564
- Grosjean, F. (2013). “Bilingualism: A short introduction,” in *The Psycholinguistics of Bilingualism*, eds F. Grosjean and P. Li (Malden, MA: John Wiley & Sons), 5–26.
- Holden, J. G. (2005). “Gauging the fractal dimension of response times from cognitive tasks,” in *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*, eds M. A. Riley and G. C. Van Orden, 268–318. Available online at: <http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp>
- Hoover, W. A., and Gough, P. B. (1990). The simple view of reading. *Read. Writ.* 2, 127–160.
- Hosp, M. K., and Fuchs, L. S. (2005). CBM as an indicator of decoding, word reading, and comprehension: do relations change with grade? *School Psychol. Rev.* 34, 9–26.
- Ihlen, E. A. F., and Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: beyond 1/f fluctuation. *J. Exp. Psychol.* 139, 436–463. doi: 10.1037/a0019098
- Ihlen, E. A. F. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. *Front. Physiol.* 3:141. doi: 10.3389/fphys.2012.00141
- Jackson, N. E., and Lu, W. H. (1992). Bilingual Precocious readers of English. *Roeper Rev.* 14, 115–119. doi: 10.1080/02783199209553404
- Jeon, E. H. (2012). Oral reading fluency in second language reading. *Read. Foreign Lang.* 24, 186–208.
- Jiang, X., Sawaki, Y., and Sabatini, J. (2012). Word reading efficiency, text reading fluency, and reading comprehension among Chinese learners of English. *Read. Psychol.* 33, 323–349. doi: 10.1080/02702711.2010.526051
- Just, M. A., Carpenter, P. A., and Woolley, J. (1982). Paradigms and processes in reading comprehension. *J. Exp. Psychol.* 111, 228–238. doi: 10.1037/0096-3445.111.2.228
- Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Phys. A* 316, 87–114. doi: 10.1016/S0378-4371(02)01383-3
- Kelty-Stephen, D. G., Palatinus, K., Saltzman, E., and Dixon, J. A. (2013). A tutorial on multifractality, cascades, and interactivity for empirical time series in ecological science. *Ecol. Psychol.* 25, 1–62. doi: 10.1080/10407413.2013.753804
- Kim, H. J. (2014). Anomalous diffusion induced by enhancement of memory. *Phys. Rev. E* 90:012103. doi: 10.1103/physreve.90.012103
- Kim, Y., and Wagner, R. K. (2015). Text (oral) reading fluency as a construct in reading development: an investigation of its mediating role for children from grades 1 to 4. *Sci. Stud. Read.* 19, 224–242. doi: 10.1080/10888438.2015.1007375
- Kim, Y., Wagner, R. K., and Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: a latent variable study of first-grade readers. *Sci. Stud. Read.* 15, 338–362. doi: 10.1080/10888438.2010.493964
- Klauda, S. L., and Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *J. Educ. Psychol.* 100, 310–321. doi: 10.1037/0022-0663.100.2.310
- Kloos, H., and Van Orden, G. C. (2010). Voluntary performance of cognitive and motor tasks. *Mind Matter* 8, 19–43.
- Kuhn, M. R., Schwanenflugel, P. J., Meisinger, E. B., Levy, B. A., and Rasinski, T. V. (2010). Aligning theory and assessment of reading fluency: automaticity, prosody, and definitions of fluency. *Read. Res. Q.* 45, 230–251. doi: 10.1598/RRQ.45.2.4
- Kuznetsov, N. A., and Wallot, S. (2011). Effects of accuracy feedback on fractal characteristics of time estimation. *Front. Neurosci.* 5:62. doi: 10.3389/fnint.2011.00062
- Lehne, M., Engel, P., Rohrmeier, M., Menninghaus, W., Jacobs, A. M., and Koelsch, S. (2015). Reading a suspenseful literary text activates brain areas related to social cognition and predictive inference. *PLoS ONE* 10:e0124550. doi: 10.1371/journal.pone.0124550
- Lems, K. (2006). “Reading fluency and comprehension in adult English language learners,” in *Fluency Instruction: Research Based Best Practices*, eds T. Rasinski, C. Blachowicz, and K. Lems (New York, NY: Guilford), 231–252.
- Lems, K. (2012). The effect of L1 orthography on the oral reading of adult English language learners. *Writ. Syst. Res.* 4, 61–71. doi: 10.1080/17586801.2011.635951



- Lesaux, N. K., Geva, E., Koda, K., Siegel, L., and Shanahan, T. (2008). "Development of literacy in second-language learners," in *Developing Reading and Writing in Second-language Learners*, eds D. August and T. Shanahan (New York, NY: Routledge), 27–60.
- Liu, H., Hu, Z., Guo, T., and Peng, D. (2010). Speaking words in two languages with one brain: neural overlap and dissociation. *Brain Res.* 1316, 75–82. doi: 10.1016/j.brainres.2009.12.030
- Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., and Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Phys. Rev. E* 66:026702. doi: 10.1103/physreve.66.026702
- Marwan, N. (2011). *Commandline Recurrence Plots*. Available online at: <http://tocsy.pik-potsdam.de/commandline-rp.php> (Accessed on August 16, 2011).
- McBride-Chang, C. (2004). *Children's Literacy Development, Texts in Developmental Psychology Series*. London: Oxford Press.
- Nelson, G. (2002). *The Singapore Corpus. International Corpus of English*. Available online at: <http://ice-corpora.net/ice/icesin.htm>
- O'Brien, B. A., and Wallot, S. (2014). "Dynamical structure of silent reading fluency in bilingual students," in *Proceedings of the XVIth European Conference on Developmental Psychology*, (Bologna: Medimondo), 175–179.
- O'Brien, B. A., Wallot, S., Haussmann, A., and Kloos, H. (2014). Using complexity metrics to assess silent reading fluency: a cross-sectional study comparing oral and silent reading. *Sci. Stud. Read.* 18, 235–254.
- Oller, D. K., Pearson, B. Z., and Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Appl. Psycholinguist.* 28, 191–230. doi: 10.1017/S0142716407070117
- Paulson, E. J. (2005). Viewing eye movements during reading through the lens of chaos theory: how reading is like the weather. *Read. Res. Q.* 40, 338–358. doi: 10.1598/RRQ.40.3.3
- Peng, C. K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time-series. *Chaos* 5, 82–87. doi: 10.1063/1.166141
- Pikulski, J. J., and Chard, D. J. (2005). Fluency: bridge between decoding and reading comprehension. *Read. Teacher* 58, 510–519. doi: 10.1598/RT.58.6.2
- Piper, B., Schroeder, L., and Trudell, B. (2016). Oral reading fluency and comprehension in Kenya: reading acquisition in a multilingual environment. *J. Res. Read.* 39, 133–152. doi: 10.1111/1467-9817.12052
- Proctor, C. P., Carlo, M., August, D., and Snow, C. (2005). Native Spanish-speaking children reading in English: toward a model of comprehension. *J. Educ. Psychol.* 97, 246–256. doi: 10.1037/0022-0663.97.2.246
- Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02
- Samuels, S. J. (2006). "Toward a model of reading fluency," in *What Research has to Say About Fluency Instruction*, eds S. Samuels and A. Farstrup (Newark, DE: International Reading Association), 24–46.
- Sandoval, T. C., Gollan, T. H., Ferreira, V. S., and Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? *The dual-task analogy. Biling. Lang. Cogn.* 13, 231–252. doi: 10.1017/S1366728909990514
- Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., and Foorman, B. R. (2004). Kindergarten prediction of reading skills: a longitudinal comparative analysis. *J. Educ. Psychol.* 96, 265–282. doi: 10.1037/0022-0663.96.2.265
- Schilling, H. E., Rayner, K., and Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. *Mem. Cogn.* 26, 1270–1281. doi: 10.3758/BF03201199
- Segalowitz, N., and Hulstijn, J. (2005). "Automaticity in bilingualism and second language learning," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (Oxford: Oxford University Press), 371–388.
- Silverman, R. D., Speece, D. L., Harring, J. R., and Ritchey, K. D. (2013). Fluency has a role in the simple view of reading. *Sci. Stud. Read.* 17, 108–133. doi: 10.1080/10888438.2011.618153
- Stephen, D. G., Boncoddio, R. A., Magnuson, J. S., and Dixon, J. A. (2009). The dynamics of insight: mathematical discovery as a phase transition. *Mem. Cogn.* 37, 1132–1149. doi: 10.3758/MC.37.8.1132
- Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (1999). *Test of Word Reading Efficiency*. Austin, TX: PRO-ED.
- Van Heuven, W. J., Dijkstra, T., and Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *J. Mem. Lang.* 39, 458–483. doi: 10.1006/jmla.1998.2584
- Van Orden, G., Kloos, H., and Wallot, S. (2011). "Living in the pink: intentionality, wellness, and complexity," in *Handbook of the Philosophy of Science, Vol. 10, Philosophy of Complex Systems*, ed C. A. Hooker (Amsterdam: Elsevier), 639–684. doi: 10.1016/b978-0-444-52076-0.50022-5
- Verhoeven, L., and van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Read. Writ.* 25, 1805–1818. doi: 10.1007/s11145-011-9346-3
- Wagner, R. K., Torgesen, J. K., and Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: PRO-ED.
- Wallot, S., Fusaroli, R., Tylén, K., and Jegindö, E. M. (2013). Using complexity metrics with RR intervals and BPM heart rate measures. *Front. Physiol.* 4:211. doi: 10.3389/fphys.2013.00211
- Wallot, S., O'Brien, B. A., Haussmann, A., Kloos, H., and Lyby, M. S. (2014). The role of reading time complexity and reading speed in text comprehension. *J. Exp. Psychol.* 40, 1745–1765. doi: 10.1037/xlm0000030
- Wallot, S., O'Brien, B., Coey, C. A., and Keltz-Stephen, D. (2015). "Power-law fluctuations in eye movements predict text comprehension during connected text reading," in *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, eds D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings et al. (Austin, TX: Cognitive Science Society), 2583–2588.
- Wallot, S., O'Brien, B. A., and Van Orden, G. (2012). "Fractal and recurrence analysis of psycholinguistic data," in *Methodological and Analytic Frontiers in Lexical Research*, eds C. Westbury, G. Jarema, and G. Libben (Amsterdam: John Benjamins), 395–430. doi: 10.1075/bct.47.18wal
- Wallot, S. (2015). From "cracking the orthographic code" to "playing with language": toward a usage-based foundation of the reading process. *Front. Psychol.* 5:891. doi: 10.3389/fpsyg.2014.00891
- Wallot, S. (2016). Understanding reading as a form of language-use: a language game hypothesis. *New Ideas Psychol.* 42, 21–28. doi: 10.1016/j.newideapsych.2015.07.006
- Webber, C. L., and Zbilut, J. P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Appl. Physiol.* 76, 965–973.
- Webber, C. L., and Zbilut, J. P. (2005). "Recurrence quantification analysis of nonlinear dynamical systems," in *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*, eds M. A. Riley, G. C. Van Orden, 26–96. Available online at: <http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jep> (Accessed on August 23, 2009).
- Wijnants, M. L., Bosman, A. M. T., Hasselman, F., Cox, R. F. A., and Van Orden, G. (2009). 1/f scaling in movement time changes with practice in precision aiming. *Nonlinear Dyn. Psychol. Life Sci.* 13, 79–98.
- Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., and Van Orden, G. (2012). An interaction-dominant perspective on reading fluency and dyslexia. *Ann. Dyslexia* 62, 100–119. doi: 10.1007/s11881-012-0067-3
- Wolf, M., and Denckla, M. B. (2004). *Rapid Automated Naming and Rapid Alternating Stimulus tests (RAN/RAS)*. Austin, TX: PRO-ED.
- Woodcock, R. W., McGrew, K. S., and Mather, N. (2007). *Woodcock-Johnson III Tests of Achievement with Normative Update*. Rolling Meadows, IL: Riverside Publishing.
- Zeno, S. M., Ivens, S. H., Millard, R. T., and Duvvuri, R. (1995). *The Educator's Word Frequency Guide*. Brewster, NY: Touchstone.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 O'Brien and Wallot. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# The Semantic Associative Ability in Preschoolers with Different Language Onset Time

Dina Di Giacomo\*, Jessica Ranieri, Eliana Donatucci, Nicoletta Caputi and Domenico Passafiume

Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy

Aim of the study is to verify the semantic associative abilities in children with different language onset times: early, typical, and delayed talkers. The study was conducted on the sample of 74 preschool children who performed a Perceptual Associative Task, in order to evaluate the ability to link concepts by four associative strategies (function, part/whole, contiguity, and superordinate strategies). The results evidenced that the children with delayed language onset performed significantly better than the children with early language production. No difference was found between typical and delayed language groups. Our results showed that the children with early language onset presented weakness in the flexibility of elaboration of the concepts. The typical and delayed language onset groups overlapped performance in the associative abilities. The time of language onset appeared to be a predictive factor in the use of semantic associative strategies; the early talkers might present a slow pattern of conceptual processing, whereas the typical and late talkers may have protective factors.

## OPEN ACCESS

### Edited by:

Simone Aparecida Capellini,  
São Paulo State University, Brazil

### Reviewed by:

Shuyan Sun,  
University of Maryland, Baltimore  
County, USA  
Michael S. Dempsey,  
Boston Medical Center, USA

### \*Correspondence:

Dina Di Giacomo  
dina.digiacomo@cc.univaq.it

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 24 May 2015

**Accepted:** 21 June 2016

**Published:** 12 July 2016

### Citation:

Di Giacomo D, Ranieri J,  
Donatucci E, Caputi N and  
Passafiume D (2016) The Semantic  
Associative Ability in Preschoolers  
with Different Language Onset Time.  
Front. Psychol. 7:1025.  
doi: 10.3389/fpsyg.2016.01025

**Keywords:** semantic associative ability, visuo-perceptive semantic, early language, delayed language, typical language

## INTRODUCTION

In the early age, children acquire the concepts observing the context and are able to organize efficiently and functionally their knowledge: progressively, the concepts develop and the semantic store emerges by using of different associative strategies. The use, recall, and functional organization of the concepts in the semantic store represent the basis of semantic competence. In that mechanism, the language represents an important cognitive factor: linguistic and conceptual development converge together in the process of early words learning (Arunachalam and Waxman, 2010). The developmental progression of knowledge is based on features of concepts: the children start from perceptual categorization to arrive to abstract categorization in order to structure the semantic store. In this process, the language represents an important increasing factor of semantic system in childhood. The language appearance in early infancy and its development represents improvement of knowledge competence (Bloom, 2000; Mandler, 2000; Booth et al., 2006; Fulkerson and Waxman, 2007; Waxman and Gelman, 2009). Arunachalam and Waxman (2010) designed mappings about the infant sensitivity to relations between words and concepts: within first year, children set words to commonalities among objects; in second year, they define precise mappings between kinds of concepts (i.e., categories of objects, properties of objects, relations among objects). Afterward, different trajectories of mappings develop: in first time mapping of nouns emerge and then the mapping for adjectives and verbs. That is due to the different informational requirement for them.

An interesting investigation is analyzing the effect of language onset time like an advantage and/or a disadvantage factor in the semantic development. Children's language emerges typically in range 12–24 months of age, but some children present a variability in terms of to begin talking: some children speak before that time and are called early talkers, whereas some else after that timing and are named late talkers.

Several studies have been conducted on the different language onset time having as focus the expressive language, the morphology and syntax, as that they represent the weaker language endowment (Rescorla, 1989, 2009; Rice et al., 2008; Rescorla and Turner, 2015). Most relevant researches have been conducted on the late talker profile identifying him as child at 2–3 years with delayed vocabulary and syntax but not significant neurological, sensory, or cognitive deficits (Desmarais et al., 2008). Moreover, Rescorla (2013) highlighted like some late talkers have expressive language delay only, whereas others have delayed receptive language.

By contrast the linguistic involvement, few studies have been focused on the effect of language onset time on the semantic competence, in particular on the use of semantic strategies basilar for the knowledge processing.

Previously, our research group investigated the semantic associative using in developmental age showing the first step of semantic processing in terms of associative strategies' using. Our findings highlighted that beginning at 4 years old, children were able to use the semantic associative relations but that competence increased during cognitive development. In particular, the ability to associate concepts using different strategies has been showed being active since the preschool age. Our research evidenced the progression of semantic associations and the roles they have in the semantic store building (Di Giacomo et al., 2012). Perceptual and then linguistic processes co-occur to develop semantic abilities. The child becomes semantically competent during preschool and early school development using sequentially perceptual and verbal encoding (Murphy, 2002; Needham et al., 2006; Coley, 2007; Nguyen, 2007; Di Giacomo et al., 2010, 2012; Herrmann et al., 2012).

Lately, we oriented our focus on the observation of the semantic strategies and the relation with the early or delayed language onset time; we have been interested to evaluate if semantic competence develops independently of language onset time, and finally, if children with early or delayed language acquisition develop semantic ability at different times; to our knowledge, few researchers have focused their interest on this topic.

Overall aim of the present study is to verify the semantic associative abilities in a preschool population tailored for different language onset time (early, delayed, and typical). We wanted to analyze if language expressive could be related to the flexibility of conceptual processing.

The study was conducted on a preschool sample from a population with language development in progress, and we assumed that the increase in linguistic competence from 3 to 6 years of age would provide a better analysis of the possible influence of language on conceptual development by visuo-perceptual elaboration.

## MATERIALS AND METHODS

### Subjects

The participants are 74 preschool children (39 female and 35 male) with mean age 4.1 years ( $SD = 0.8$ ) distributed in three groups defined by phase of language onset: (i) the Early Language (EL) group included 17 children with mean age 3.9 ( $SD = 0.8$ ) with early language onset (Mean = 7.8 months and  $SD = 0.5$ ); (ii) the Typical Language Language (TL) group included 39 with mean age 4.4 ( $SD = 0.8$ ) and with typical language onset (mean = 11.3 months and  $SD = 1.2$ ); (iii) the Delayed Language group (DL) included 18 children with mean age 3.8 ( $SD = 0.7$ ) with delayed language onset (mean = 17.3 months and  $SD = 2.9$ ). The distribution of the sample in the three groups was made on the basis of pediatric evaluations, parents' reports on the basis of Rescorla's criteria (Rescorla, 1989): age of acquisition of first words, age of gesture indication, and age of spontaneous use of first phrases (Table 1).

Excluded children have been n.74 because their performance have been under the threshold by Raven test (see Test).

All children lived with both parents.

### Test

A standardized psychological battery was administered.

*Raven's Colored Progressive Matrices* (Italian Adaptation Belacchi et al., 2008) is a non-verbal test widely applied in the evaluation of general intelligence, and is composed of 36 items. The subject was asked to choose from a set of six, the piece that was missing in a target pattern. The standard score was analyzed. The Raven's Colored Progressive Matrices was used to measure the cognitive competence of the subjects in order to exclude those with cognitive deficits/difficulties.

*Prova di Associazione Semantica* (PAS, Semantic Associative Task, Di Giacomo and Passafiume, 2014) is a visuo-perceptual task to evaluate the semantic associative abilities. It was carried out on native Italian speaking children. The task was composed of two sets: Naming and Matching tasks.

### Naming Task

The Naming task consists of 40 drawing items representing objects applied in the Matching task. The examiner asks the subject to say the name of the drawn object (Figure 1). The Naming task is a preliminary test to measure the children's ability to recognize the targets used in the Matching task (cut-off is 75% correct responses). The score is the sum of correct responses.

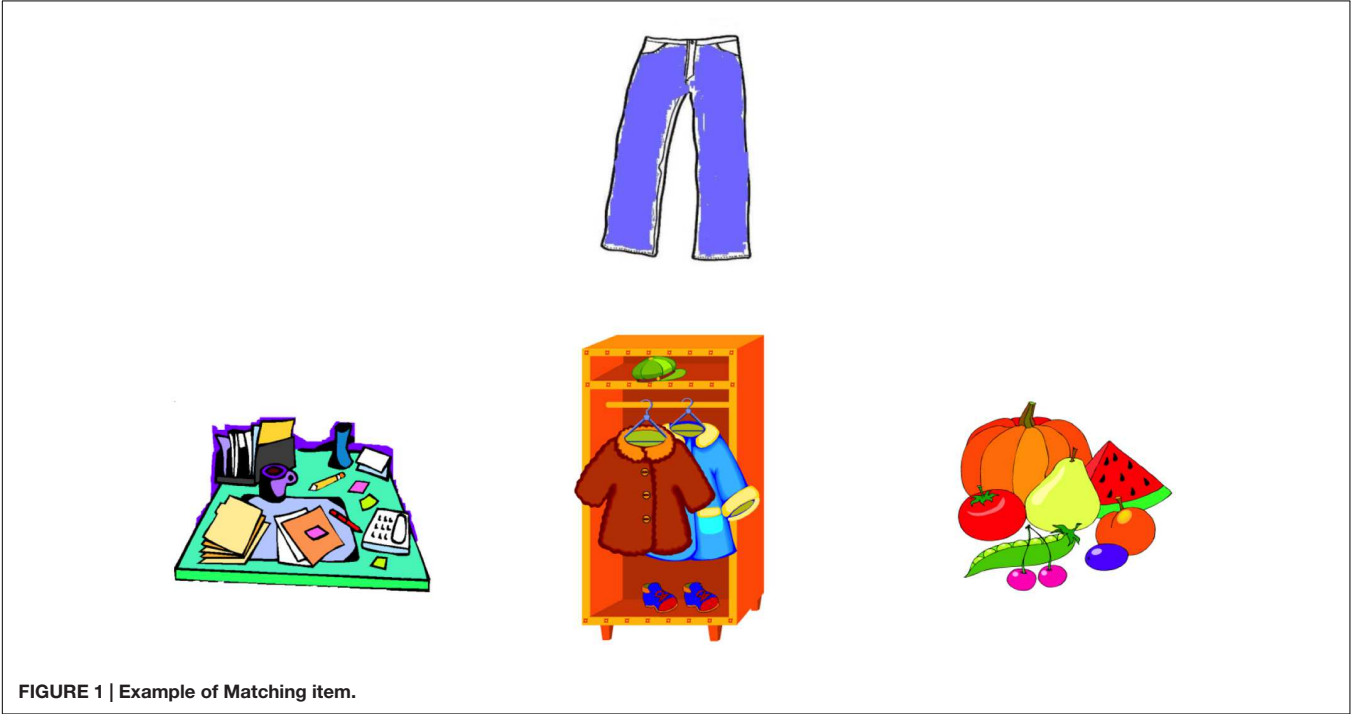
### Matching task

Matching task is composed of 40 items and each item includes one target object [and three other objects (see Figure 2)]. The examiner asks to the subject to indicate which one of the three choices (objects) is related better than others to the target. The items investigate four semantic associative relations: (i) Function, (ii) Part/Whole, (iii) Contiguous, and (iv) Superordinate). The associative relations were as follows: the Function category consists of pairing an object with its use (e.g., scissors and to cut); the Part/Whole category consists of pairing an object with its single part (e.g., fish and fin); the Contiguous category

TABLE 1 | Demographic data of the participants.

Variables	Sample			F	p
	Early language	Typical language	Delayed language		
N° (Total n = 74)	17	39	18		
Male	10	14	11		
Female	7	25	7		
Age	3.9 (±0.8)	4.3 (±0.7)	3.8 (±0.7)		
Age mother	37.9 (±5.6)	36.7 (±5.5)	39.7 (±5.4)		
Age father	40.8 (±5.4)	40.5 (±6.3)	42.3 (±2.6)		
Gesture age	7.5 (±1.15)	8.2 (±1.7)	12.1 (±3.2)	32.4	0.00*
First word age	9.2 (±2.5)	10.9 (±1.9)	15.6 (±4.2)	18.2	0.00*
Raven test	14.2 (±4.4)	17.2 (±4.9)	14.2 (±5.3)	2.4	0.09
t Raven test	5.4 (±3.7)	4.8 (±1.7)	3.4 (±1.5)	2.6	0.08

\*statistically significant.



consists of pairing an object with its complement (e.g., pencil and eraser); the Superordinate category consists of pairing an object with its class membership (e.g., dog and animal). Three trial items applied. The score was the sum of the correct responses. The Cronbach  $\alpha$  value is: function = 0.83; part/whole = 0.86; contiguity = 0.80; superordinate = 0.80).

In addition, the time was measured for the subject's completion of the Naming and Matching tasks.

Procedure

The children have been recruited in pediatric ambulatory and kindergarten school. The children have been evaluated by Psychologists in individual sessions lasting 45 min in a quiet and dedicated room. The scoring of psychological tests was get by judges were blind by the study's objectives. Parents have been

proposed a individual interview lasted at least 1 h in order to have more information about linguistic ability of their children. Written informed consensus by parents was mandatory and obtaneid.

Data was inserted in the Case Report form builded for this research.

Ethic Statement

The study was carried out with the Positive Opinion of Ethic Commetee of University of L'Aquila (Italy).

Plan Statistical Analysis

The data were submitted to statistical analysis with value  $\alpha < 0.05$ . The statistical analysis were performed through the Statistica software.

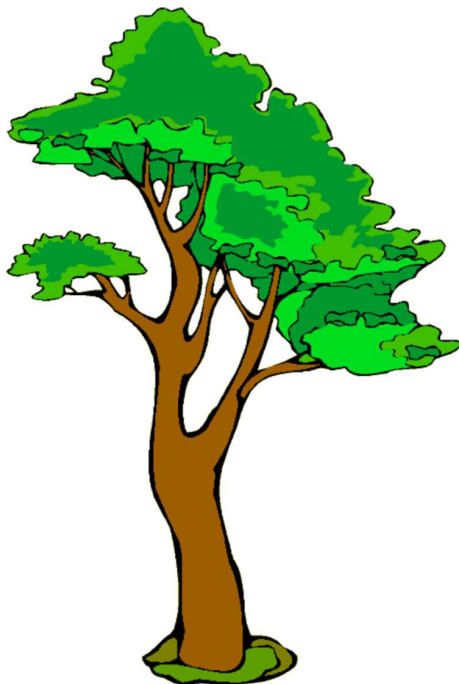


FIGURE 2 | Example of Naming item.

TABLE 2 | Raw scores of age groups in the experimental tasks.

Tests	Age groups		
	3-years-old	4-years-old	5-years-old
Naming	29.6 ( $\pm 7.0$ )	34.4 ( $\pm 3.7$ )	34.9 ( $\pm 3.2$ )
Matching	30.5 ( $\pm 5.8$ )	37.9 ( $\pm 1.1$ )	34.4 ( $\pm 6.9$ )
Function	7.9 ( $\pm 1.5$ )	9.1 ( $\pm 0.8$ )	9.6 ( $\pm 0.6$ )
Part/Whole	7.2 ( $\pm 1.5$ )	8.9 ( $\pm 1.5$ )	9.0 ( $\pm 0.8$ )
Contiguity	7.9 ( $\pm 1.4$ )	9.0 ( $\pm 0.9$ )	9.6 ( $\pm 0.6$ )
Superordinate	7.3 ( $\pm 2.3$ )	9.2 ( $\pm 0.9$ )	9.6 ( $\pm 0.7$ )

Descriptive statistics (mean and standard deviation for numeric variables, frequencies, frequencies for categorical variables) were processed for all variables examined.

An ANOVA analysis was applied to match the semantic performance difference in three groups (TL, EL, and DL), and then we conducted the *post hoc* analysis (Tukey test). Suddenly, we conducted MANOVA to compare the age groups and the language onset time groups to evaluate the effect of aging and the language onset time on the semantic performance. The aging effect is expected.

## RESULTS

Aim of the research was to analyze the semantic associative performance in early developmental age. Our focus has been the use of associative strategies in the range age 3–6 years old, in a tailored sample by different language onset time.

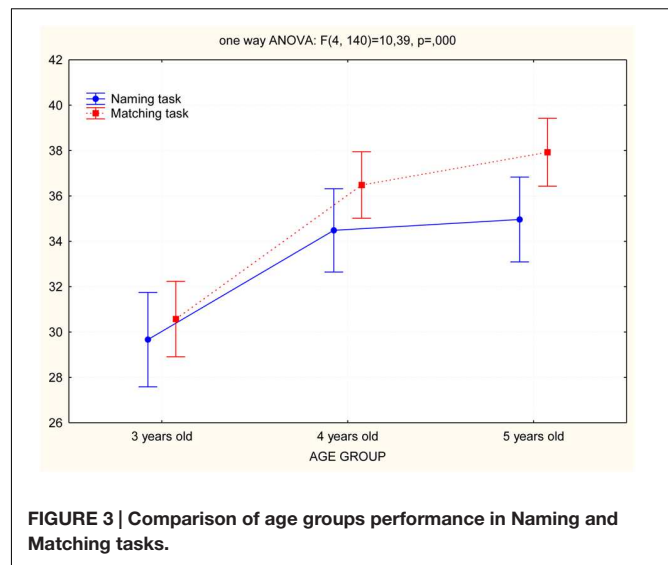
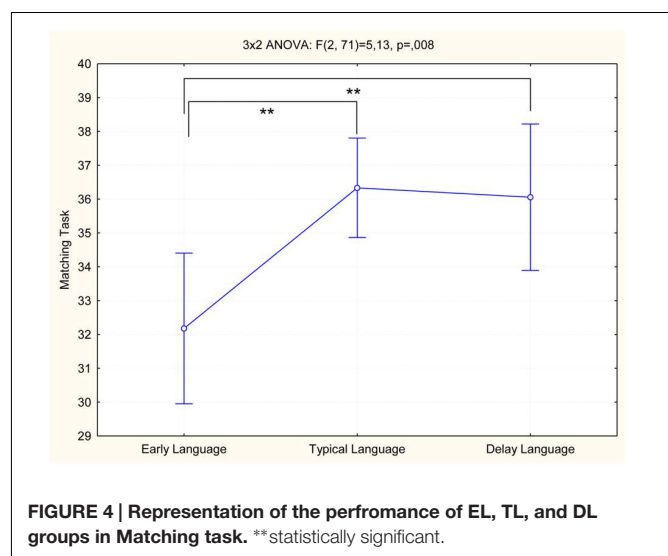


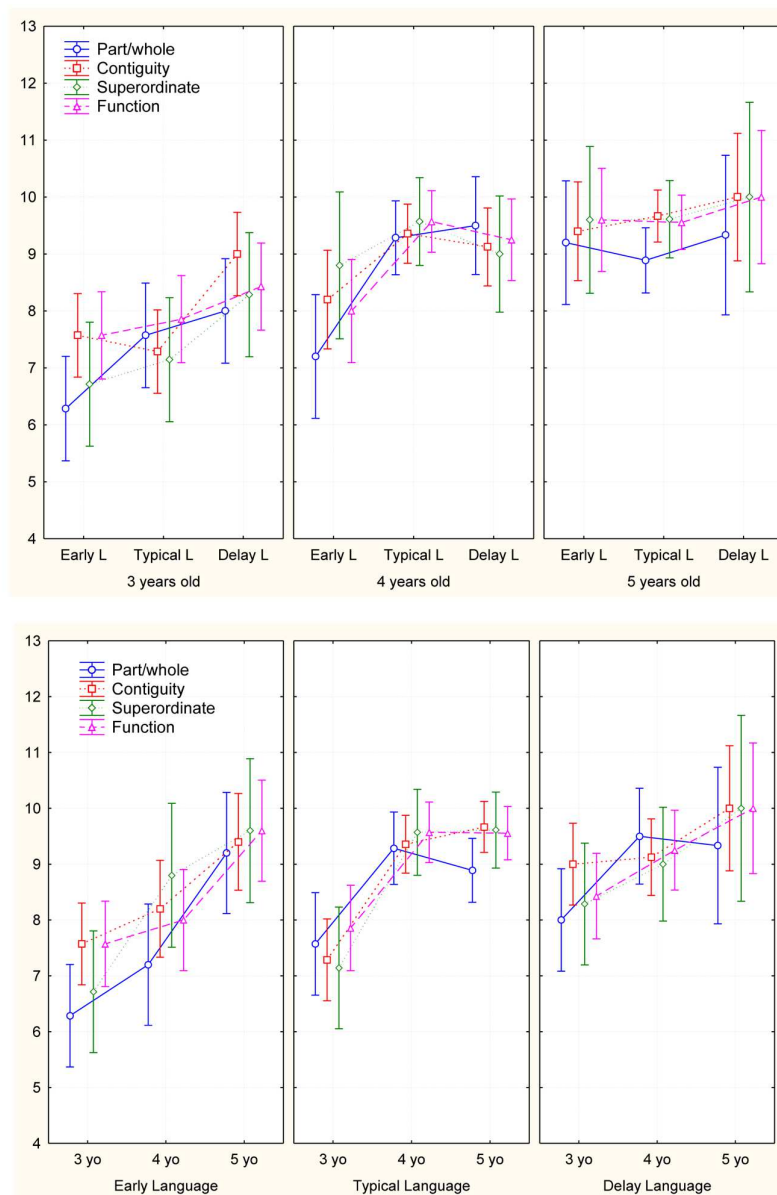
TABLE 3 | Raw scores of language onset time groups in the experimental tasks

Tests	Language onset time groups		
	Early language	Typical language	Delayed language
Naming	31.9 ( $\pm 3.9$ )	34.4 ( $\pm 4.5$ )	33.7 ( $\pm 6.3$ )
Matching	32.5 ( $\pm 5.7$ )	36.0 ( $\pm 4.7$ )	34.4 ( $\pm 6.9$ )
Function	8.3 ( $\pm 1.3$ )	9.2 ( $\pm 1.1$ )	8.7 ( $\pm 1.7$ )
Part/Whole	7.5 ( $\pm 1.8$ )	8.7 ( $\pm 1.2$ )	8.6 ( $\pm 1.5$ )
Contiguity	8.3 ( $\pm 1.2$ )	9.0 ( $\pm 1.3$ )	8.5 ( $\pm 2.2$ )
Superordinate	8.2 ( $\pm 2.1$ )	8.9 ( $\pm 1.7$ )	8.4 ( $\pm 2.2$ )



First, we wanted to analyze the influence of age in the elaboration of semantic associative strategies. The sample has been divided in three groups by the chronological age: (i) 3-year-old group was composed of 21 subjects, (ii) 4-year-old group was





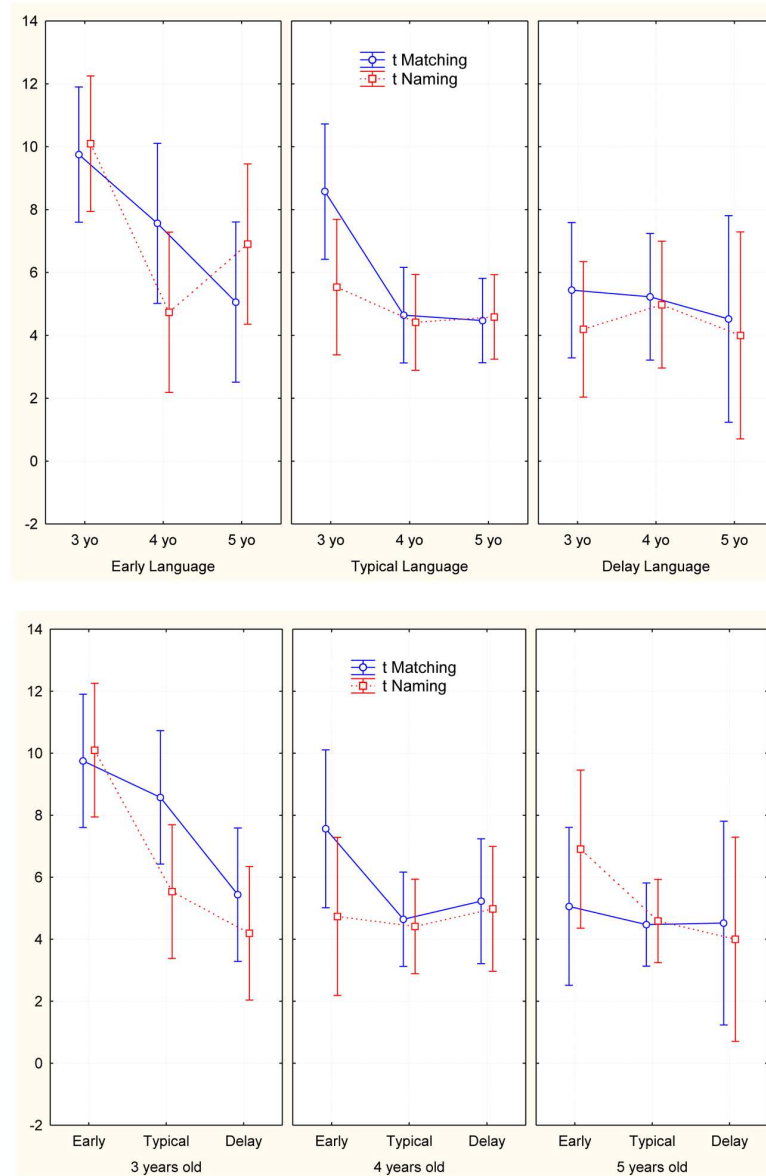
**FIGURE 5 | Representation of the Matching task performance by age and language onset groups.**

composed of 26 subjects, and (c) 5-year-old group was composed of 27 subjects. In **Table 2**, we reported raw score.

A MANOVA 3 (age groups)  $\times$  2 (tasks: Naming, Matching) evidenced significant difference among the three groups in the two tasks [Naming:  $F(2,71) = 8.4$ ;  $p = 0.001$ , and  $\eta^2 = 0.19$ ; Matching:  $F(2,71) = 23.5$ ;  $p < 0.0001$ , and  $\eta^2 = 0.39$ ]. The *Post hoc* analysis (Tukey test) showed that in the Naming task, the 3-year-old group was significantly different from the 4-year-old ( $p < 0.002$ ), and 5-year-old groups ( $p < 0.001$ ) while no significant difference were found between the 4- and 5-year-old groups. Significant differences were also found in the Matching task: the 3-year-old group

was less able than the 4-year-old ( $p < 0.001$ ) and 5-year-old groups ( $p < 0.004$ ; **Figure 3**). The expected results have confirmed out the previous data (Di Giacomo et al., 2012).

Then, we have conducted a statistical analysis to evaluate the performance of the three language onset time groups (EL, TL, and DL) in the associative test (Naming and Matching task). **Table 3** reported the raw score of the sample distributed in language onset time. A  $3 \times 2$  MANOVA showed differences between language onset time groups in semantic tasks [ $F(4,140) = 2.94$ ;  $p = 0.02$ , and  $\eta^2 = 0.78$ ]. The *Post hoc* analysis (Tukey test) evidenced different performance between language onset groups only in



**FIGURE 6 | Representation of time execution in Naming and Matching tasks.**

the Matching task: the EL group's scores were lower than TL ( $p < 0.001$ ) and DL ( $p < 0.004$ ) groups (**Figure 4**).

Besides, a 3 (language onset time groups)  $\times$  3 (age groups)  $\times$  4 (types of semantic associations: function, part/whole, contiguity, and superordinate) MANOVA showed a significant difference among the age groups [ $F(8,124) = 2.3$ ;  $p < 0.001$ , and  $\eta^2 = 0.25$ ] and the onset language groups [ $F(8,124) = 5.34$ ;  $p < 0.02$ , and  $\eta^2 = 0.13$ ], but no significant interaction between age and language onset time groups. This result is interesting: the aging effect isn't affect the semantic associative performance of children with different language onset time (**Figure 5**).

Finally, we have analyzed the execution time ( $t$ ) of sample in Naming and Matching tasks. A 3 (EL, TL, and DL groups)  $\times$  3 (age groups)  $\times$  2 ( $t$  Naming and Matching tasks) MANOVA evidenced significant differences in language onset time groups [ $F(4,128) = 2.7$ ;  $p < 0.03$ , and  $\eta^2 = 0.78$ ] and age groups [ $F(4,128) = 3.4$ ;  $p < 0.01$ , and  $\eta^2 = 0.09$ ]; Tukey test showed in EL performance resulting slower than TL and DL groups in Matching task ( $p < 0.001$ ); TL and DL groups performance appear similar. The *post hoc* on age groups performance evidenced the older children (4- and 5-year-olds) faster than younger (3-year-olds) ( $t$  Naming:  $p < 0.05$ ,  $t$  Matching:  $p < 0.008$ ) (**Figure 6**).

## DISCUSSION AND CONCLUSION

The present study proposed to analyze the impact of the language onset time in the development of associative strategies using. Particularly, we wanted to verify if the semantic ability in early childhood could be affected by language onset time, reflecting specific features as well as linguistic competence.

Our data showed that language onset time does not seem to affect directly the use of examined semantic associative abilities. The children improve their using of associative strategies during cognitive development, without significant linkage to verbal production. The data evidenced that the children with delayed language are able to use the associative strategies as well as the children with typical language: these performance appear in the elaboration of information and in execution time of semantic task.

Our findings showed developing semantic ability isn't related primarily to the language onset time. The performance of Delayed Language Onset Group didn't be different from the Typical Language Onset Group on the Matching Task; moreover, DL performance have differed from EL in both measurements (correctness and execution time). The early language children have been less efficient than the other subjects of two groups in the concepts association and the use of single associative relations. The early language group appeared weak in the use of contiguity and part/whole relations.

Our results suggest that semantic association competence and the age of linguistic production aren't directly linked, even though the early word production could predict a weakness in the managing of the linkage of the concepts; in contrary, the delayed linguistic production didn't seem to influence the development of associative strategies.

Several studies demonstrated the delayed lexical activation could reflect a weakness in language development and favoring bloomer and/or late talker outcomes (Rescorla, 2005; Rice et al., 2008). Rice et al. (2008) conducted a follow up study of the evolution of the performance of late talking children at 3-year-old: the research demonstrated the persistence of linguistic impairment connected to the syntactic and grammatical deficit and a relative deficit in the semantic quotient (verbal task) with important involvement of spontaneous language.

Rescorla (2009) showed that the linguistic difficulties persisted into adolescence. Follow-up studies evidenced such as the language that initially have evolution difficulties during the develop maintain critical even if supported linguistic rehabilitation interventions. These conclusions are supported by several reports (Rescorla, 2005; Rice et al., 2008). Our research

highlighted the importance of the strength of conceptual flexibility in subjects with delayed language onset.

Few studies have focused on the early talkers. Our results suggested that the early talkers have a weakness in their semantic competence: though their verbal production is early, their development of conceptual associative strategies is later than typical and delayed talkers. They performed well in the Naming task, but not in the Matching Task. We added to a development model of the semantic and conceptual stores the finding that late talkers demonstrate stronger conceptual processing. Several studies investigated the grammatical and lexical difficulties in cognitive development; in our research the late talkers showed more competence in the use of semantic associative strategies. Furthermore, early talking can be considerate a predictive factor for use in educational systems to improve semantic ability since to early time. In the applied psychology, and in particular in the educational stimulation, our results can contribute to the formulation of intervention planning more efficiently focused on integration of the competences of the conceptual and semantic memory on child with delayed onset language. Semantic categorization can be used as a competence on which building, through use of plans to stimulate and promote linguistic performance. The meanings of the words and the linkages between them might improve the outcomes of the educational stimulation, and later, increase verbal production.

## AUTHOR CONTRIBUTIONS

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

## FUNDING

This study was conducted by Italian National Grant awarded to DG – Call 2008 'Future in Research'(MIUR) (code RBFR08A5NE).

## ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance of the International Neuropsychological Society's Research Editing and Consulting Program, particularly Dr. Carol Armstrong. for assistance for English language editing.

## REFERENCES

- Arunachalam, S., and Waxman, S. R. (2010). Language and conceptual development. *Cogn. Sci.* 1, 548–558. doi: 10.1002/wcs.37
- Belacchi, C., Scalisi, T. G., Cannoni, E., and Cornoldi, C. (2008). *Manuale CPM Coloured Progressive Matrices: Standardizzazione Italiana*. Firenze: Giunti O.S. Organizzazioni Speciali.
- Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.
- Booth, A. E., Waxman, S. R., and Huang, Y. T. (2006). Conceptual information permeates word learning in infancy. *Dev. Psychol.* 41, 491–505. doi: 10.1037/0012-1649.41.3.491
- Coley, J. (2007). The human animal: developmental changes in judgments of taxonomic and psychological similarity among humans and other animals. *Cogn. Brain Behav.* 11, 733–756.
- Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., and Roileau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *Int. J. Lang. Commun. Disord.* 43, 361–389. doi: 10.1080/13682820701546854

- Di Giacomo, D., De Federicis, L. S., Pistelli, M., Fiorenzi, D., and Passafiume, D. (2012). Semantic associative relations and conceptual processing. *Cogn. Process.* 13, 55–62. doi: 10.1007/s10339-011-0399-7
- Di Giacomo, D., and Passafiume, D. (2014). *Prova di Associazione Semantica*. Milano: Franco Angeli.
- Di Giacomo, D., Pistelli, M., Fiorenzi, D., and Passafiume, D. (2010). Costruzione del network semantico nell'infanzia: ruolo svolto dal sistema di codifica percettivo. *G. Psicol.* 4, 71–78.
- Fulkerson, A. L., and Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: evidence from 6- and 12-month-old. *Cognition* 105, 218–228. doi: 10.1016/j.cognition.2006.09.005
- Herrmann, P., Medin, D., and Waxman, S. (2012). When humans become animals: development of the animal category in early childhood. *Cognition* 122, 74–79. doi: 10.1016/j.cognition.2011.08.011
- Mandler, J. M. (2000). Perceptual and conceptual processes in infancy. *J. Cogn. Sci.* 1, 3–36. doi: 10.1207/s15327647jcd0101n\_2
- Murphy, G. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.
- Needham, A., Cantlon, J. F., and Ormsbee Holley, S. M. (2006). Infants' use of category knowledge and object attributes when segregating objects at 8.5 months of age. *Cogn. Psychol.* 53, 345–360. doi: 10.1016/j.cogpsych.2006.05.003
- Nguyen, S. (2007). Cross-classification and category representation in children's concepts. *Dev. Psychol.* 43, 719–731. doi: 10.1037/0012-1649.43.3.719
- Rescorla L. (2005). Age 13 language and reading outcomes in late-talking toddlers. *J. Speech Lang. Hear. Res.* 48, 459–472. doi: 10.1044/1092-4388(2005/031)
- Rescorla, L. (1989). The language development survey: a screening tool for delayed language in toddlers. *J. Speech Hear. Disord.* 54, 587–599. doi: 10.1044/jshd.5404.587
- Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: support for a dimensional perspective on language delayed. *J. Speech Lang. Hear. Res.* 52, 17–30. doi: 10.1044/1092-4388(2008/07-0171)
- Rescorla, L. (2013). "Late-talking toddlers: a fifteen-year follow-up," in *Late Talkers Language Development, Intervention, and Outcomes*, eds L. A. Rescorla and P. S. Dale (Baltimore, MD: Brookes), 219–240.
- Rescorla, L., and Turner, H. (2015). Morphology and syntax in late talkers at 5 age. *J. Speech Lang. Hear. Res.* 58, 434–444. doi: 10.1044/2015\_JSLHR-L-14-0042
- Rice, M., Taylor, C., and Zubrick, S. (2008). Language Outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *J. Speech Lang. Hear. Res.* 51, 394–407. doi: 10.1044/1092-4388(2008/029)
- Waxman, S. R., and Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends Cogn. Sci.* 13, 258–263. doi: 10.1016/j.tics.2009.03.006

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Di Giacomo, Ranieri, Donatucci, Caputi and Passafiume. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Reading Comprehension Assessment through Retelling: Performance Profiles of Children with Dyslexia and Language-Based Learning Disability

Adriana de S. B. Kida<sup>1\*</sup>, Clara R. B. de Ávila<sup>2</sup> and Simone A. Capellini<sup>1</sup>

<sup>1</sup> Department of Speech-Language Pathology and Audiology, Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, Brazil, <sup>2</sup> Department of Speech-Language Pathology and Audiology, Universidade Federal de São Paulo, São Paulo, Brazil

## OPEN ACCESS

### Edited by:

Douglas Kauffman,  
Boston University School of Medicine,  
USA

### Reviewed by:

Ana Miranda,  
Universidad de Valencia, Spain  
Jaclyn M. Dynia,  
Ohio State University, USA

### \*Correspondence:

Adriana de S. B. Kida  
adrianabatista@gmail.com

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 01 December 2015

**Accepted:** 10 May 2016

**Published:** 01 June 2016

### Citation:

Kida ASB, Ávila CRB and Capellini SA  
(2016) Reading Comprehension  
Assessment through Retelling:  
Performance Profiles of Children with  
Dyslexia and Language-Based  
Learning Disability.  
Front. Psychol. 7:787.  
doi: 10.3389/fpsyg.2016.00787

**Purpose:** To study reading comprehension performance profiles of children with dyslexia as well as language-based learning disability (LBLD) by means of retelling tasks.

**Method:** One hundred and five children from 2nd to 5th grades of elementary school were gathered into six groups: Dyslexia group (D;  $n = 19$ ), language-based learning disability group (LBLD;  $n = 16$ ); their respective control groups paired according to different variables – age, gender, grade and school system (public or private; D-control and LBLD-control); and other control groups paired according to different reading accuracy (D-accuracy; LBLD-accuracy). All of the children read an expository text and orally retold the story as they understood it. The analysis quantified propositions (main ideas and details) and retold links. A retelling reference standard (3–0) was also established from the best to the worst performance. We compared both clinical groups (D and LBLD) with their respective control groups by means of Mann–Whitney tests.

**Results:** D showed the same total of propositions, links and reference standards as D-control, but performed better than D-accuracy in macro structural (total of links) and super structural (retelling reference standard) measures. Results suggest that dyslexic children are able to use their linguistic competence and their own background knowledge to minimize the effects of their decoding deficit, especially at the highest text processing levels. LBLD performed worse than LBLD-control in all of the retelling measures and LBLD showed worse performance than LBLD-accuracy in the total retold links and retelling reference standard.

Those results suggest that both decoding and linguistic difficulties affect reading comprehension. Moreover, the linguistic deficits presented by LBLD students do not allow these pupils to perform as competently in terms of text comprehension as the children with dyslexia do. Thus, failure in the macro and super-structural information processing of the expository text were evidenced.

**Conclusion:** Each clinical group showed a different retelling profile. Such findings support the view that there are differences between these two clinical populations in the non-phonological dimensions of language.

**Keywords:** reading comprehension, retelling, simple view of reading, dyslexia, recall pattern

## INTRODUCTION

Reading comprehension assessment by retelling a previously read text enables direct access to the expression of mental representation built by the reader (Leslie and Caldwell, 2009; Reed and Vaughn, 2012) without any interference or facilitation. It also evidences one's competence to both identify relevant information of the previously read text and integrate these ideas into a cohesive and coherent global scheme (Orrantia et al., 1990; León, 1991; García Madruga et al., 1996). Retelling, allows for an outlining of different levels of comprehension, these being the reading processing product resultant from the micro, macro and super-structures (Squires et al., 2014; Kida et al., 2015). The identification of the total retold ideas is a direct measure of how the reader operates with the basic units of the text – the propositions – and it reflects his ability to keep them in mind (Orrantia et al., 1990; Carlisle, 1999). These abilities are intricately connected to the micro-structural comprehension, therefore, to the local information processing level (Kintsch and Keenan, 1973; Frazier and Fodor, 1978). The total retold ideas, mainly when considering the importance of each one to the textual chain (main ideas or details), gives clues of the reader's abilities to choose relevant information from the text and to start the global processing of the text macro-structure. After choosing, generalizing and leaving ideas out, skills involved in this process, the reader starts establishing connections between each piece of information, in other words, the links. Retelling also provides the product of global processing at a macro-structural level. Such product may be assessed by the total links made between the retold ideas, and by the retelling reference standard (Squires et al., 2014). The measurement provided by the retelling standard takes into account the analysis of the set of retold links considering its relevance for the text central chain. For that reason, such measurement helps determine the comprehension level achieved by the reader and, therefore, the way he reaches the text super-structure. The reference standard of the retelling of expository texts evaluates the way the reader organizes his ideas toward an established central goal (Bustos Ibarra, 2009; Squires et al., 2014), guided by his knowledge of the text structure (Richgels et al., 1987; Roller, 1990).

Hence, the set of retelling measurements – total retold ideas, total links and the retelling reference standard – may provide hints of how the reader builds his understanding of the text and at what level of processing difficulties remain when reading comprehension does not take place (Owens et al., 1979; Orrantia et al., 1990; Bernhardt, 1991).

Reading comprehension difficulties result from varied types of reading deficits, identified by the reader's performance in automatic recognition of written words and/or in oral comprehension. Three clinical groups are known: (1) Readers with specific reading comprehension deficits; (2) Readers with specific decoding deficits – dyslexia; (3) Readers with both deficits – known as language-based learning disability (LBD) deficits (Catts et al., 2003) or also as mixed deficits (Catts et al., 2005b; Cain and Oakhill, 2006). The differences between the latter two clinical groups are primarily in non-phonological

language dimensions. Children with LBD show, besides a phonological processing impairment, a typical dyslexia symptom, significant deficits of oral comprehension (Aaron, 1991; Catts and Kahmi, 2005) with hindrance to vocabulary, morph syntax, and text structural processing, even when their non-verbal abilities are preserved. These linguistic deficits are, therefore, broad and, they interact directly with reading competences, resulting in different manifestations, thus making reading problems in the LBD group more evident and, equally broad (Catts, 1993; Catts et al., 1997, 2005a; Bishop and Snowling, 2004).

Although children with dyslexia and with LBD knowingly show difficulties of very different nature, it is acknowledged that reading comprehension can be impaired in both cases. Children with dyslexia may present reading comprehension difficulties influenced by their decoding deficits, despite their good oral comprehension. Their slow and inaccurate word recognition may limit sentence and text processing speed, thus resulting in comprehension impairments (LaBerge and Samuels, 1974; Perfetti, 1985; Shankweiler et al., 1999). Pupils with learning disabilities, in turn, show deficit in reading comprehension as a consequence of poor decoding abilities and of more general language deficit (Stanovich and Siegel, 1994; Ellis et al., 1996; Aaron et al., 1999).

Literature has not yet shown if retelling allows for an identification of different performance profiles in reading comprehension among the different cases of reading impairment. The presence of heterogeneous groups (poor readers, reading disabled children, and children with learning disabilities) in the searched studies does not help understand the effects of the deficits in different competences (decoding and language) upon reading comprehension.

Studies showed that these children with learning difficulties, designated to the sample upon their teachers' recommendation or chosen according to the identification of deficits in their reading performance on specific evaluation tests, significantly retold less pieces of information. Furthermore, they showed worse oral discourse management of the retelling due to their difficulties in adequately choosing main ideas rather than details (Williams, 1991; Curran et al., 1996; Carlisle, 1999; Reed and Vaughn, 2012). Even in studies that encouraged retelling, requesting children to add more pieces of information at the end of their narrative, showed that there was no retelling expansion. The retelling of children with academic learning difficulties showed no expansion of number of ideas (Bridge and Tierney, 1981; Zinar, 1990; Reed and Vaughn, 2012), as opposed to the performance of good readers.

When considering such particular nature of deficits that affect children with dyslexia and children with LBD, it is expected the investigated clinical groups (D, LBD) to show distinct performance profiles in reading comprehension at different levels of text processing. The chances of the retelling identify the quality of mental representation generated by the text, by means of measuring the total of retold ideas (main and details), links and the reference standard that each one of those clinical groups grasped, would help understand which strategies of base-text construction could be impaired. Then, it is necessary to compare the performance of each clinical group with the performance of

its pairs of the same age, gender, school grade, as well as with pupils of the same reading accuracy level.

Cain et al. (2000) proposed an experimental design based on the use of comparison by pairing levels of competence. The idea of this design is to compare an interest group with pairs of the same age and school grade as a means of controlling the effect of variables, such as language development and schooling, upon reading performance. Moreover, comparing accuracy competence groups helps not only to demonstrate if the performance of a clinical group is below its pairs of the same level of development and schooling, but also to advance in understanding the cause of reading difficulties (Frith and Snowling, 1983; Siegel and Ryan, 1988). The comparison between comprehension competence of a clinical group and younger pupils of equivalent decoding competence may help determine what is the most probable explanation for the relationship between both competences: is reading comprehension proficiency due to linguistic development or to decoding competence? If the group without complaints of reading difficulties, paired by age, gender and schooling shows better reading comprehension competence, we may suppose that decoding competence may interfere in comprehension access, once this competence distinguishes these groups. However, if the performance of pupils with difficulties is better than pupils paired by level of accuracy, for example, we may presume that the schooling and/or the language development, a distinguishing factor between the groups, may favor comprehension performance. At last, if the performance of pupils with difficulties is worse than the group paired by level of accuracy, it is presumable that, once decoding is controlled, the deficits in language competence may interfere in the performance of pupils with difficulties.

This experimental design is intended to help clarify the nature of reading comprehension difficulties faced by children with dyslexia and by children with language-based learning disabilities. Findings are intended to help in the comprehension of the necessary supports and to facilitate the planning of required intervention by each of these groups of readers.

## Purpose

This study aimed at characterizing oral retelling profiles made by children with dyslexia and LBLD, after reading an expository text.

First, we intended to identify the points of comprehension breakdown at different text processing levels (micro structure, macro structure, and super structure), as well as to measure the effects of deficits based on the expected performance according to age and schooling. For such, clinical groups were compared with its controls of the same level of development and schooling.

Afterward, we meant to understand those comprehension breakdowns (micro-structure, macro-structure, and super-structure) based on the language competence shown by the clinical groups. For such, the decoding variable was controlled through pair ups according to accuracy level.

Thus, we tried to answer the following research questions starting from the designed hypotheses.

- (a) Do children with dyslexia, who present restricted decoding difficulties, necessarily show reading comprehension impairments? At what text processing level do decoding difficulties interfere to the point of impairing reading comprehension? Can skilful language competence compensate for decoding deficits? Also, at which level processing of textual information could language favor reading comprehension?

In the light of those questions, we adopted the following hypotheses:

- Children with dyslexia (D) will show reading comprehension difficulties when compared with the group paired by two development parameters – age and schooling (school grade and school system) – (D-control). By that, it may be possible to assume that the presented decoding deficits may generate negative effects on reading comprehension.
  - Children with dyslexia will show better performance when compared with the reading accuracy level group (D-accuracy). By that, it may be possible to assume that children with dyslexia are able to employ their most skilful language competences and to make use of their knowledge, built throughout their longer schooling period, in favor of reading comprehension performance. For that reason, they will show better chances of processing information than younger children with the same decoding competences.
- (b) Will children with LBLD, who show both decoding and language competence difficulties, present specific alterations at certain levels of reading comprehension, or a broader compromised profile in processing textual information? At which levels of processing textual information may the deficits present in language interfere in reading comprehension activities?

In turn, for children with LBLD, we proposed the following hypotheses:

- Children with LBLD will show difficulties in every level of processing of textual information when compared with the children of the same level of development, gender and schooling (LBLD-control).
- Children with LBLD will show similar or worse performance when compared with pupils of the same reading accuracy level (LBLD-accuracy). That result will allow us to assume that their language competences will not be enough to compensate for their decoding difficulties, making their reading comprehension performance similar or even worse than this control group. Such result will depend on the extent of the effect of linguistic competences this population has.

## MATERIALS AND METHODS

The sample consisted of 105 students, native Brazilian Portuguese speakers from the 2nd to the 5th grades of Elementary

School, without complaints nor indicators of hearing or visual impairment, neurological, behavioral or cognitive disability. They comprised six groups: (a) D: 19 children (53% male, average age: 127 months,  $SD = 16.1$ ) with clinical developmental dyslexia diagnostic; (b) D-control: 19 children (53% male, average age: 123 months,  $SD = 15$ ) without complaints of reading difficulties. This group was paired up with D according to age, gender, school system and schooling parameters; (c) D-accuracy: 19 children (47% male, average age: 123 months,  $SD = 2.1$ ) paired up with D according to reading accuracy; (d) LBLD: 16 children (81% male, average age: 122.5 months,  $SD = 14.7$ ) with clinical diagnostic of LBLD; (e) LBLD-control: 16 children (81% male, average age: 121 months,  $SD = 10.4$ ). This group was paired with LBLD according to age, gender, school system and schooling; (f): LBLD-accuracy: 16 children (44% male, average age: 86.5 months,  $SD = 3.9$ ) paired with LBLD according to reading accuracy.

Clinical group participants (D and LBLD) were recruited through cross-disciplinary diagnostics (neurologist, neuropsychologist, psycho-pedagogue, and speech therapist) carried out at Laboratório de Investigação dos Desvios de Aprendizagem do Centro de Estudos da Educação e da Saúde da Faculdade de Filosofia e Ciências – CEES/FFC/UNESP Marília – SP (Learning Deviation Investigative Laboratory of the Centre of Health and Education Studies of the Philosophy and Science College – CEES/FFC/UNESP Marília – SP) and at Laboratório dos Desvios de Aprendizagem do Hospital das Clínicas da Faculdade de Medicina – HC/FM/UNESP – Botucatu – SP (Learning Deviation Laboratory of the Clinics Hospital of Medicine College – HC/FM/UNESP – Botucatu – SP).

D pupils showed (1): an expected intelligence quotient (I.Q. equal to or higher than 80) in psychological evaluation, (2) the presence of significant discrepancy between the verbal and the execution quotients, with differences leaning toward the execution IQ, with lowering score in the digit subtest and good performance at vocabulary and arithmetic subtests taken according to the expected values for the age at WISC-III (Wechsler, 2002); (3) low performance in the reading of both the isolated word task, according to parameters established for the standard test for the Brazilian school population (Stein, 1994) and the pseudo words (Arduini et al., 2006; Salgado and Capellini, 2008); (4) performance impairment of phonological short-term memory, according to the expected schooling parameters (Kessler, 1997; Tabakim, 2008); (5) poor performance in the rapid serial naming task (Denckla and Rudel, 1974) according to parameters of the Brazilian school population (Simões, 2006); (6) performance impairment in the phonological awareness task, showing a lowering of more than 1.5 dp of the total score for the schooling (Capovilla and Capovilla, 1998).

Language-based learning disability participants met the following inclusion criteria (Puranik et al., 2006): (1) history of previous language impairment or academic difficulties in early school years; (2) intellectual quotient below average (minimum I.Q. of 80 points) with the absence of discrepancy between the verbal and the execution intellectual quotient at the WISC-III psychological assessment (Wechsler, 2002); (3) the same or better performance than the percentile 25 (below average level) at the Raven's Progressive Matrices, with schooling parameters

taken into account (Raven et al., 1988); (4) good performance at the Wisconsin Card Sorting test classification, with schooling parameters taken into account (Heaton et al., 2005); (5) poor performance in the tasks of reading isolated words, according to the Brazilian schooling population parameters for writing and Arithmetic (Stein, 1994).

The control group participants were recruited in Elementary Schools of the same city. Besides meeting the recruitment criteria established for the whole sample, these children did not show history of speaking nor language impairment, of academic or reading difficulties, neither suggestive signals of sensory alterations, neurological and cognitive impairment, according to their teacher's designation.

The study of pairing up clinical groups (D and LBLD) with their controls of the same age, gender, schooling and school system (D-control and LBLD-control) in relation to the age variable was carried out by means of the one-way analysis of variance (ANOVA One Way), making use of age (in months) as a dependent variable. Bonferroni Tests were performed in order to verify the existence of differences between the pairs of groups. Results showed that there were no differences between the groups [D and D-control:  $F(1,36) = 0.46$ ,  $p = 0.500$ ,  $\eta^2 = 0.907$ ; LBLD and LBLD-control:  $F(2,35) = 6.38$ ,  $p = 0.500$ ,  $\eta^2 = 0.026$ ].

Mann-Whitney test was used to assess the pairing up of clinical groups (D and LBLD) with their controls according to age and schooling (D-control and LBLD-control) and with their controls according to reading accuracy (D-accuracy and LBLD) for the decoding variable. Rate (number of words read per minute) and accuracy (number of correct words read per minute) were used as dependent variables and group as a fixed factor.

The comparison of the decoding variables (rate and accuracy) resultant from the single item task (Pinheiro, 2011) showed that the clinical groups (D and LBLD) presented lower figures when compared with their control pairs according to age, gender, and schooling (Table 1). In turn, the comparison between the clinical groups and their controls, paired according to their level of reading, showed similar figures (Table 2). Results proved the decision of pairing up these groups appropriate, considering the adoption of the experimental design.

## Procedures

### Protocol of Retelling after Reading

Four expository texts were carefully written about subjects that were not part of the private and public school programs, neither in their previous grades nor in the intended evaluated grade. Such criterion aimed at minimizing the effects of the participants' previous knowledge of the reading comprehension assessment.

The texts proved appropriate for each school grade. A previous study revealed that the texts were appropriate for: readability (analyzed through the Flesch Index), syntax complexity (Indexes: number of words and sentences of the texts, number of sentences within paragraphs, occurrence of content words, pronouns per syntagma, and number of linkers) and vocabulary complexity (Type/Token Index), attested parameters that interfere in comprehension (Aluísio et al., 2008; McNamara et al., 2010; McNamara et al., 2012). All of these measurements were achieved



**TABLE 1 | Study of the sample pair up based on reading fluency variables – Comparison between clinical groups and control groups according to age, gender and schooling.**

Variables		Groups			
		D (N = 19)	D-control (N = 19)	LBLD (N = 16)	LBLD-control (N = 16)
Reading rate	Mean	25.9	56.5	23,1	63,1
	SD	11.2	11.3	8,4	19,1
	U	9.50		11.00	
	P-value	$p < 0.001$		$p < 0.001$	
	Results	D < D-control		LBLD < LBLD-control	
Reading accuracy	Mean	11.9	49.2	8,5	54,6
	SD	9.1	12.2	5,5	6,3
	Z	9.00		6.50	
	P-value	$p < 0.001$		$p < 0.001$	
	Results	D < D-control		LBLD < LBLD-control	

Mann–Whitney test. D, Dyslexia Group; D-control, Dyslexia control group by age, gender, and grade; LBLD, Language-based learning disability group; LBLD-control, Language-based learning disability control group by age, gender and grade. Significant at  $p < 0.05$ .

**TABLE 2 | Pairing up data between clinical groups and their control groups paired according to reading accuracy upon reading fluency variables.**

Variable		Groups			
		D (N = 19)	D-accuracy (N = 19)	LBLD (N = 16)	LBLD-accuracy (N = 16)
Rate	Mean	25.9	24.0	23.1	27.6
	SD	11.2	8.12	8,4	15.2
	U	153.50		119.50	
	P-value	0.219		0.234	
	Results	D = D-accuracy		LBLD = LBLD-accuracy	
Accuracy	Mean	11.9	11.6	8.7	9.1
	SD	9.1	10.3	5.5	6.3
	U	170.5		124.50	
	P-value	0.625		0.237	
	Results	D = D-accuracy		LBLD = LBLD-accuracy	

Mann–Whitney test. D, Dyslexia Group; D-accuracy, Dyslexia control group by reading accuracy; LBLD, Language-based learning disability group; LBLD-accuracy, Language-based learning disability control group by reading accuracy. Significant at  $p < 0.05$ .

through the *CohMetrix-Port* computerized tool (Scarton and Aluísio, 2010) and were compared with the expected parameters for texts of each researched school grade (Kida, 2015). These reference figures were established based on the analysis of 15 schoolbook collections of the Portuguese Language (total of 918 texts), approved by the Plano Nacional do Livro Didático (Schoolbook National Plan) – PNLD 2013 (Brasil, 2012), intended to the teaching of pupils from the 2nd to the 5th grades of Elementary School. Changes were made in the texts once comparisons showed inadequacies on a certain assessed parameter. Such changes assured that the final texts had the desirable syntax complexity and readability for each school grade.

The analysis of the retellings was carefully sifted through three assessors based on the propositions of each text. Such analyses were compared and the propositions classified, under consensus, as main ideas and details. For such, the importance of each

proposition for the main chain of the text and the nature of the conveyed information (explicit or implicit) were considered. The identified propositions composed the screening and allowed for the identification of total retold ideas, a parameter used to assess the processing of text micro-structure (Sánchez, 2002; Bustos Ibarra, 2009).

Assessors also identified the existing links in the text (Sánchez, 2002; Bustos Ibarra, 2009). Links were understood as causal connection between main ideas, between main ideas and details or between details. Every identified link composed the screening and allowed for the assessment of the quality of the global processing of textual macro-structure.

Another adopted parameter was the retelling standard reference, established according to the following criteria (Coté et al., 1998; Gonzales, 2008; Bustos Ibarra, 2009): Standard 3–presence of all the links between main ideas, together with, at least, one link between main ideas and details; Standard

2–presence of links between main ideas, without the presence of links between main ideas and details; Standard 1–presence of, at least, one link, no matter its classification; Standard 0–absence of causal connections of any type. These standards were established to reflect different ways of organizing memorized ideas toward the central objective of the text, thus allowing for the observation of the super-structural processing of the text by the reader (Bustos Ibarra, 2009).

### Data Collection Procedures

The assessed children were instructed to read the text that was assigned for their school grade the habitual way they read for comprehension (oral or silently). No time limit was defined. They retold the previously read information when they considered they were ready. Their retellings were recorded for later transcription and analysis.

### Procedures of Data Analysis

#### *Procedures of analysis of inter-rater agreement*

Four speech therapists were trained to analyze the retellings. After training, transcriptions of 200 retellings (50 of each school grade) from previous researches were used and helped achieve the inter-rater agreement. Such measure intended to guarantee reliability. To estimate the Reliability Indexes, the Kappa inter-observer agreement measurement was used, establishing 0.70 as a minimum level (Urbina, 2007). The resulting indexes for each one of the texts are in **Table 3**.

#### *Procedures of retelling analysis*

Once the inter-observer agreement was proven, the participants' retelling transcriptions were distributed among the four coders. It was a blind assessment as for age, gender, and schooling of participants, as well as their groups.

Coders identified each one of the ideas and retold links and scored one point for each piece of information identified. Then, they summed the total ideas and links. Eventually, they identified the retelling standard with its respective scorings.

Gross results of total ideas and links were converted based on z-scores, as lower, average, and upper levels, to which 0, 1, and 2 scores were, respectively, assigned (Kida, 2015). The data resulting from the analysis of each retelling were tabulated and submitted to statistical analysis.

### Data Analyses

The normality test indicated the presence of non-normal distribution for the analysis variables. The comparative analysis of performance of clinical groups with their respective controls – comparison of independent groups, two by two – was carried out through the Mann–Whitney test for the total ideas, total links and retelling standard. The significance level adopted was  $p < 0.05$ .

The extent of the effect size for the Mann–Whitney test was assessed by approximation of the distributions of test statistics for the Z distribution, once this is a non-parametric test. Thus, the following calculation was made:  $r = Z/\sqrt{N}$ . The analysis criteria for effect size (proposed upon the Cohen  $r$ ) and adopted for this study were: great effect = 0.5; average effect = 0.3; low effect = 0.1 (Fritz et al., 2012).

The statistical pack IBM SPSS Statistics – version 22 (pt) – was used in the analysis mentioned above.

## RESULTS

### Study Results of the Retelling Profile of Children with Developmental Dyslexia

**Table 4** shows that children with dyslexia (D) presented performance similar to their pairs of the same age, gender and schooling (D-control) for the variables of total ideas and of retold links, as well as of retelling standard after reading.

The comparative investigation of total retold ideas showed that D retold less central ideas, significantly differing from its pairs of the same age and schooling (D: 0.42/D-control: 1.11;  $U = 150.00$ ,  $p = 0.004$ ,  $r = 0.542$ , I.C. 95%: lower limit =  $-0.7$ , upper limit =  $-1.09$ ). However, D and D-control performances were similar when compared with the total retold details (D: 0.58/D-control = 0.89;  $U = 21.00$ ,  $p = 0.154$ ).

Comparisons between children with dyslexia (D) and their pairs of the same level of reading accuracy (D-accuracy) showed that children with dyslexia presented better performance evidenced by the greater number of retold links ( $r = -0.1286$ , I.C. 95%: lower limit =  $-0.905$ , upper limit =  $-0.1667$ ) and by the retelling score ( $r = 0.1286$ , I.C. 95%: lower limit =  $-0.905$ , upper limit =  $-0.1664$ ), as observable in **Table 5**.

The comparative investigation of the total retold ideas showed that dyslexic children (D) presented similar performance to the observed in children with the same level of accuracy (D-accuracy; Main ideas – D: 0.42/D-accuracy: 0.11;  $p = 0.096$ ; Details: D: 0.58/D-accuracy: 0.26;  $p = 0.096$ ).

### Study Results of the Retelling Profile of Children Diagnosed with Language-Based Learning Disability (LBLD)

**Table 6** shows the results of the comparison between the performance of children with LBLD and their pairs of the same age, gender, and schooling (LBLD-control) as for the total ideas, links and retelling standard. LBLD showed significantly poorer performance than LBLD-control in every measure of retelling (Total retold ideas:  $r = 0.1286$ , I.C. 95%: lower limit =  $-0.0872$ , upper limit =  $0.1770$ ; Total retold links:  $r = 0.0952$ , I.C. 95%: lower limit =  $-0.0431$ , upper limit =  $0.1473$ ; RRS:  $r = 0.0168$ , I.C. 95%: lower limit =  $-0.0076$ , upper limit =  $0.0262$ ). The comparative analysis showed that children from the LBLD group presented similar performance both in main ideas and details (Main ideas: LBLD: 0.25/LBLD-control: 1.06,  $U = 44.00$ ,  $p = 0.268$ ,  $r = 0.0169$ , I.C. 95%: lower limit =  $-0.0713$ , upper limit =  $0.1050$ ; Details: LBLD: 0.25/LBLD-control: 0.87,  $U = 44.00$ ,  $p = 0.236$ ,  $r = 0.0233$ , I.C. 95%: lower limit =  $-0.0102$ , upper limit =  $0.0364$ ).

The performance of children with LBLD and of their pairs of the same level of accuracy (LBLD-accuracy) showed to be similar in every variable of retelling performance, as observable in **Table 7**.

**TABLE 3 | Inter-rater Agreement index for variables of analysis of retelling after reading expository texts assigned to each school grade.**

Retelling analysis variables	Expository texts			
	2nd grade	3rd grade	4th grade	5th grade
Tot_ID	0.730	0.817	0.704	0.703
Tot_Links	0.928	0.898	0.752	0.801
RRS	0.862	0.961	0.760	0.850

*Kappa Test. Tot\_ID, total of ideas retold; Tot\_links, total of links retold; RRS, retelling reference standard. Cut-Off: 0.70.*

**TABLE 4 | Reading comprehension performance of D and D-control in retelling task.**

		Group		Mann-Whitney test	Result
		D	D-control		
Tot_ID	Mean	0.63	1.00	$p = 0.354$	D = D-control
	SD	0.495	0.577		
Tot_Links	Mean	0.263	0.895	$p = 0.402$	D = D-control
	SD	0.452	0.809		
RRS	Mean	0.263	0.895	$p = 0.258$	D = D-control
	SD	0.452	0.567		

*Tot\_ID, total of ideas retold; Tot\_links, total of links retold; RRS, retelling reference standard; SD, standard deviation; D, Dyslexia Group; D-control, Control group by age, sex and grade.*

**TABLE 5 | Reading comprehension performance of D and D-accuracy in retelling task.**

		Group		Mann-Whitney test	Result
		D	D-accuracy		
Tot_ID	Mean	0.632	0.315	$p = 0.488$	D = D-accuracy
	SD	0.495	0.477		
Tot_Links	Mean	0.263	0.105	$p = 0.005^{***}$	D > D-accuracy
	SD	0.452	0.315		
RRS	Mean	0.263	0.105	$p = 0.005^{***}$	D > D-accuracy
	SD	0.452	0.3153		

*Tot\_ID, total of ideas retold; Tot\_links, total of links retold; RRS, retelling reference standard; SD, standard deviation; D, Dyslexia Group; D-accuracy, control group by reading accuracy; significant at  $p < 0.05$ , \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .*

**TABLE 6 | Reading comprehension performance of LBLD and LBLD-control in retelling task.**

		Group		Mann-Whitney test	Result
		LBLD	LBLD-control		
Tot_ID	Mean	0.437	1.187	$p = 0.002^{***}$	LBLD < LBLD-control
	SD	0.512	0.403		
Tot_Links	Mean	0.125	1.125	$p = 0.003^{***}$	LBLD < LBLD-control
	SD	0.341	0.806		
RRS	Mean	0.062	0.875	$p = 0.047^{**}$	LBLD < LBLD-control
	SD	0.250	0.719		

*Tot\_ID, total of ideas retold; Tot\_links, total of links retold; RRS, retelling reference standard; SD, standard deviation; LBLD, Language-based learning disability group; LBLD-control, control group by age, sex, and grade; significant at  $p < 0.05$ ; \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .*

As for the retold ideas, however, the typical children paired according to level of accuracy (LBLD-accuracy) retold as much main ideas and details as the children with LBLD (Main ideas: LBLD: 0.25/LBLD-accuracy: 0.56,  $U = -16.00$ ,  $p = 0.346$ ; details: LBLD: 0.25/ LBLD-Ac: 0.75,  $U = -44.00$ ,  $p = 0.579$ ).

## DISCUSSION

### Profile of Retelling of Children with Developmental Dyslexia

The comparison between the performances of children with developmental dyslexia and typical children of the same age, gender, and schooling showed that decoding problems did not interfere in the performance measured by the most general parameters of the retelling (total ideas and retold links, reference standard). These results oppose the initial hypotheses.

Puranik et al. (2008) also observed similar performance within pupils with developmental dyslexia and their controls of the same age and schooling assessed through a written retelling task. Both groups showed the same total retold ideas. The authors assign the good performance of the dyslexia group to the children's skilful language competences, which allowed them to compensate for their reading decoding deficits.

However, even considering the fact that the total retold ideas did not affect group performances at the present study, children with important reading decoding impairment showed poorer performance as for the number of main ideas retold. Such result shows that children with dyslexia were less efficient than typical readers of the same age and schooling in identifying and choosing main ideas, an important competence for textual processing at its macro-structural level. These children's worse performance in retelling main ideas suggests that decoding difficulty interfere in the use of macro-rules employed in the construction of the mental representation of the text (Kintsch, 1988, 1998), causing loss of recent information as well as difficulty in leaving out less relevant data (Weaver and Dickinson, 1979).

Similar results were found by Snyder and Downey (1991) and Nascimento et al. (2011). These finding showed that retelling after reading with a smaller number of main ideas

happened among readers with poor performance in reading fluency, evidenced by the low rate and accuracy estimates in tasks of reading recognition of isolated words. However, in this study, decoding automaticity failures and the effects they implied to the processing of syntactic information impaired comprehension of the connections of ideas (Snyder and Downey, 1991; Nascimento et al., 2011). The absence of readiness for information processing at a micro-structural level did not allow poor readers to draw their attention and their meta-cognitive resources to process high-level information (identification of the main subject of the text and recognition of the textual structure), primal in regulating macro-rule employment (Orrantia et al., 1990; Snyder and Downey, 1991; García Madruga et al., 1996; Nascimento et al., 2011). Inefficient use of macro-rules (choosing and leaving information out based on its importance for the chain of the text) determined the effects on the progressive construction of the mental representation during reading.

Overall, literature implies that competition between the decoding and the comprehension competences may impair access to the meaning of words and to the quality of syntactic processing (Vogel, 1975) and, consequently, interfere in the construction of a mental representation and/or the transference of information to the long-term memory (Shankweiler and Liberman, 1972; Juel, 1988; Shankweiler et al., 1999; Bowey, 2000). Furthermore, studies suggest that the natural limitation of the operational memory to manage information during reading at the presence of decoding problems compromises reading comprehension. Competition between decoding and comprehension by the operational memory builds a barrier that can restrict the use of language high-level processing systems, required by the global comprehension of the previously read content (Perfetti and Lesgold, 1977; Snyder and Downey, 1991).

Data analysis of the present study did not reveal, as expected, a worse performance of the group of children with developmental dyslexia for macro and super-structural measurements, in other words, for total links and retelling standard.

That way, although the decoding effects on reading comprehension are known, a theoretical view seeks to explain how underlying competence differences may influence reading

**TABLE 7 | Reading comprehension performance of LBLD and LBLD-accuracy in retelling task.**

		Group		Mann-Whitney test	Result
		LBLD	LBLD-accuracy		
Tot_ID	Mean	0.437	0.625	$p = 0.239$	LBLD = LBLD-accuracy
	SD	0.512	0.500		
Tot_links	Mean	0.125	0.375	$p < 0.001$	LBLD < LBLD-accuracy
	SD	0.341	0.619		
RRS	Mean	0.006	0.187	$p < 0.001$	LBLD < LBLD-accuracy
	SD	0.250	0.403		

*Tot\_ID*, total of ideas retold; *Tot\_links*, total of links retold; *RRS*, retelling reference standard; *SD*, standard deviation; *LBLD*, Language-based learning disability group; *LBLD-accuracy*, control group by reading accuracy; significant at  $p < 0.05$ .



at an interactive and compensatory perspective (Stanovich, 1980; Perfetti and Roth, 1981). Evidences that the nature of the parallel processing of reading may compensate for decoding difficulties have been broadly demonstrated. Children with dyslexia show better competence at using contextual facilitation than typical readers, because of the adequacy of their oral comprehension competence (Nation and Snowling, 1988). Attention to contextual information within texts frequently serve children with dyslexia to solve decoding ambiguities based on the collection of hints and the use of their previous knowledge, always through the action of their preserved cognitive and linguistic competences (Nation and Snowling, 1988).

A study carried out by Shankweiler et al. (1999) showed that pupils with reading decoding difficulties can compensate such difficulties with the use of their good linguistic competences through a “top-down” processing. But, contradicting the results of the present study Shankweiler et al. (1999) indicate that this compensatory competence is limited, viewing that it was not able to level the performance of dyslexic pupils to the one of typical readers. Lots of authors argue that the competence of compensating decoding impairments is more often observable in students with dyslexia who have been studying for a longer period of time, such as adolescents and adults (Campbell and Butterworth, 1985; Simmons and Singleton, 2000). At a younger age, compensation seems to contribute only to literal processing of information, not reaching inferential competences (Miller-Shaul, 2005).

Although compensation may be a hypothesis, data observation allows us to suppose that the similar performance of pupils with dyslexia and their pairs of the same age and schooling happened because of the low performance of the control group. When considering the control of the school system variable (private or public), impairment becomes evident as for the performance of reading comprehension also of readers taken as proficient by their teachers.

National assessment of reading comprehension in Brazilian Primary Schools and Middle School/Junior High (from the 1st to the 9th grades) shows that 45% of students with 4 years of schooling, after the beginning of the literacy process, present low reading comprehension levels (Brasil, 2014). Although they manage to deal with explicit information, to make connections between the text information or, to a certain extent, make use of their knowledge of the world, most of these students can only use these competences in simple texts (Bridon and Neitzel, 2014), below the expected level for their schooling.

Data collected from the control of the decoding effects on reading comprehension showed that the performance of children with dyslexia was significantly better, considering the number of retold links and the retelling standard achieved when compared with its pairs of the same level of accuracy. Data suggest that better language competence and the experience reached through a longer schooling period of children with dyslexia provide them with greater competence of connecting the processed ideas at a macro-structural

level, as well as of incorporating them into a broader textual scheme.

These findings corroborate the hypothesis that decoding difficulties may be minimized when two linguistic competences are present. Thus, it may be assumed that the best language competences of children with dyslexia can be used in favor of a more efficient processing of the macro-structure of the text, having, therefore, greater competence in integrating main identified ideas (total links) and in its integration into a general textual scheme (reference standard of the retelling).

Orrantia et al. (1990) defend the idea that competent readers use varied cognitive operations to achieve global meaning (Meyer, 1984; Kintsch, 1998): choice, generalization, integration, and suppression of propositions. However, when these strategies come together with recognition of the global structure of the text, organization and integration of the propositions chosen at a coherent global scheme are even more efficient. These high-level competences are connected both to a good language development, for integration and choice, and to the reading experience, for recognition of the global structure of the text (García Madruga et al., 1996). That way, although children with dyslexia presented the same possibilities of identifying and retelling main ideas than pupils of the same level of reading, their greater reading experience has probably allowed them to transfer such gains to the processing of certain macro and super-structural levels of the text, results corroborated by Weaver and Dickinson (1979) and Kornev and Balciuniene (2014).

## Profile of the Retelling of Children Diagnosed with Language-Based Learning Disability

Children with LBLD showed important difficulties in reading comprehension, presenting poorer performance at all text processing levels when compared with their pairs of the same age and schooling. This group's language difficulties were also a key factor for its comprehension performance to level to younger children with the same reading accuracy for text processing at its macro and super-structural levels.

These results confirm that LBLD suffer the effects of decoding deficits and it does not show the same compensation competence evidenced in the performance of children with dyslexia, once they present important deficits in linguistic abilities and essential competences for reading comprehension. These findings are possibly explained by language deficits, which prevent efficient activation of mechanisms implicated in reading.

Studies carried out with pupils with LBLD found less retold main ideas when compared with pupils of the same age and schooling (Williams, 1991; Curran et al., 1996; Carlisle, 1999; Puranik et al., 2008). Under control of the vocabulary variable, Carlisle (1999) also demonstrated that pupils with learning disabilities showed to be less able to understand and use textual structure as support to integrate the most important ideas of the text. Hansen (1978) points out that these pupils tended to present greater number of intrusions, in other words, they presented a greater number of pieces of information that did not belong to the text, a frequent behavior among children

with important comprehension difficulties. Also, literature meta-analysis indicates that lots of studies report procedures that encourage pupils with learning disabilities to complement their retellings and, on the contrary, this strategy does not result in improvement of the total retold ideas or even the establishment of connections between the pieces of information (Reed and Vaughn, 2012). All of these findings suggest that the presence of language deficits restrains text processing at macro and super-structural levels, required for a global comprehension of what was read.

One of the explanations about comprehension difficulties presented by pupils with LBLD goes beyond decoding interference. It considers the influence of deficits in integrating information based on syntax problems, resulting from difficulties expressed throughout language development. Among the observed interferences, difficulties to accomplish inferences such as anaphora (Oakhill et al., 1986) and the worst performance of referential continuity in stories (Garnham et al., 1982) must also be highlighted. These manifestations would cause problems in understanding the role of linking elements and the organization of pieces of information – conjunctions and discourse markers – and would make the text processing at its macro and super-structural levels difficult, expressed by restricted number of retold links and by the worst reference standard of the retelling.

It is important to point out that, among limitations of the present, the low effect size found does not allow for the generalization of findings presented so far. New studies must be carried out in order to broaden evidences found in the present research.

## CONCLUSION

The reading comprehension assessment through a task of oral retelling after reading indicated that children with dyslexia and with LBLD showed difficulty in making sense of a read expository text. However, the groups presented impairments at different levels of text processing and different coverage of reading comprehension deficits.

Children diagnosed with developmental dyslexia showed more restricted impairments at macro-structural levels, considering the lower efficiency demonstrated to identify and choose main ideas when compared with typical pupils of the same age and schooling.

Children diagnosed with LBLD showed broader difficulties, impairing every level of text processing, in other words, micro, macro and super-structural levels.

The comparison of clinical groups with the performance of their typical pairs of the same reading accuracy also confirms the existence of differences in performance profile of dyslexic children and of children with LBLD. Such difference is possibly due to the distinct conditions of development of non-phonological dimensions of language observed in clinical groups.

Children diagnosed with developmental dyslexia showed better competence in retelling links between ideas present in

the previously read text, and also achieved better retelling standards than pupils of the same decoding level. These findings suggest that the language competences and the knowledge acquired throughout schooling provide these children with better abilities to connect processed ideas at a macro-structural level, as well as to incorporate them into a broader textual scheme. Pupils with LBLD showed greater difficulty to connect ideas and also to build a global representation of the text.

At last, the different performance profiles in reading comprehension identified in the investigation of clinical groups with different types of Reading Disabilities suggest the possibility of achieving important indicators by means of the retelling after reading task. These indicators may help in a more precise diagnostic of reading comprehension impairment, making it more precise and specific. The study also shows the viability of using the retelling protocol as a means of accessing the mental elaboration built during reading, in which it is possible to determine the points of breakdown that may compromise reading comprehension. Its directive analysis may also be fundamental for clinical use, since, the precise identification of the comprehension level where difficulties remain allows for a more specific intervention upon these deficits. Besides, better understanding of difficulties may bring important educational developments as it enables the adoption of facilitating strategies and more assertive adaptations.

## ETHICS STATEMENT

This study was approved by the Research Ethics Committee of the Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP – Marília (No. 0928/2014). The assessments started after the following: 1) authorization to collect data at the Clinic of Learning Disabilities at the Clinical Hospital of Faculdade de Medicina de Botucatu – Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP – Botucatu and the Laboratory for Investigation of Learning Disabilities at the School of Philosophy and Sciences, Universidade Estadual Paulista “Júlio de Mesquita Filho” – CEES/FFC/UNESP – Marília (SP).

## AUTHOR CONTRIBUTIONS

AK prepared the project and the evaluation tool, collected and analyzed the statistics of the survey data, identified the literature, and wrote the article. CÂ collaborated in drawing up the assessment tool, in the discussion of survey data, and the preparation of the article. SC supervised the study, participated in the discussion of the data, and the preparation of the article.

## FUNDING

This project was supported by National Counsel of Technology and Scientific Development (CNPq- Brasil).

## ACKNOWLEDGMENTS

The authors want to thank Gabriela Bueno, Suelen Rossi, Mariana Oliveira, Adriana Martins, Esmeralda Damasceno, Thiago Takahashi, and Anderson Nascimento for their valuable help in conducting the analysis of retellings. The authors also

want to thank Patrícia Lúcio, Hugo Cogo-Moreira, Carolina Carvalho, Paola Okuda, and Vânia Lima for their helpful suggestions and comments on earlier versions of this paper. We appreciate the thoughtful revisions and helpful suggestion made by the associate editor and the reviewers, which clearly contributed to substantially improving the paper.

## REFERENCES

- Aaron, P. G. (1991). Can reading disabilities be diagnosed without using intelligence tests? *J. Learn. Disabil.* 24, 178–186. doi: 10.1177/002221949102400306
- Aaron, P. G., Joshi, M., and Williams, K. A. (1999). Not all reading disabilities are alike. *J. Learn. Disabil.* 32, 120–137. doi: 10.1177/002221949903200203
- Aluísio, S. M., Specia, L., Pardo, T. A. S., Maziero, E. G., and Fortes, R. P. M. (2008). “Towards brazilian portuguese automatic text simplification systems,” in *Proceedings of the 8th ACM Symposium on Document Engineering*, São Paulo, 16–19.
- Arduini, R., Capellini, S. A., and Ciasca, S. M. (2006). Comparative study neuropsychological and neuroimaging evaluations in children with dyslexia. *Arq. Neuropsiquiatr.* 64, 369–375. doi: 10.1590/S0004-282X2006000300004
- Bernhardt, E. B. (1991). *Reading Development in a Second-Language*. Norwood, NJ: Ablex.
- Bishop, D. V. M., and Snowling, M. (2004). Developmental dyslexia and specific language impairment: same or different? *Psychol. Bull.* 2, 201–212.
- Bowey, J. A. (2000). Recent developments in language acquisition and reading research: the phonological basics in children’s reading difficulties. *Aust. Educ. Dev. Psychol.* 17, 5–31.
- Brasil (2012). *Guia de Livros Didáticos: PNLD 2013: Letramento e Alfabetização e Língua Portuguesa*. Brasília, DF: Ministério da Educação, Secretaria de Educação Básica.
- Brasil (2014). *SAEB/Prova Brasil 2011 - Primeiros Resultados*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Ministério da Educação, Diretoria de Avaliação da Educação Básica.
- Bridge, C. A., and Tierney, R. J. (1981). The inferential operations of children across text with narrative and expository tendencies. *J. Read. Behav.* 13, 201–214.
- Bridon, J., and Neitzel, A. A. (2014). Competências leitoras no Saeb: qualidade da leitura na educação básica. *Educ. Real.* 39, 437–462. doi: 10.1590/S2175-62362014000200006
- Bustos Ibarra, A. (2009). *La Competencia Retórica y el Aprendizaje de la Lengua Escrita. ¿Se Puede Hablar de una Competencia Específica?* doctoral tesis. Universidad de Salamanca, Salamanca.
- Cain, K., and Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *Br. J. Educ. Psychol.* 76, 683–696. doi: 10.1348/000709905X67610
- Cain, K., Oakhill, J., and Bryant, P. (2000). Investigating the causes of reading comprehension failure: the comprehension-age match design. *Read. Writ.* 12, 31–40. doi: 10.1023/A:1008058319399
- Campbell, R., and Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject: a developmental case with associated deficits in phonemic processing and awareness. *Q. J. Exp. Psychol.* 37, 375–396. doi: 10.1080/14640748508400944
- Capovilla, A. G. S., and Capovilla, F. C. (1998). Prova de consciência fonológica: desenvolvimento de dez habilidades da pré-escola à segunda série. *Temas Desenvol.* 7, 14–20.
- Carlisle, J. F. (1999). Free recall as a test of reading comprehension for students with learning disabilities. *Learn Disabil. Q.* 22, 11–22. doi: 10.2307/1511148
- Catts, H., Fey, M., and Tomblin, B. (1997). Language basis of reading disabilities. *Paper Present at Society for the Scientific Study of Reading*, Chicago, IL.
- Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *J. Speech Lang. Hear Res.* 36, 948–958. doi: 10.1044/jshr.3605.948
- Catts, H. W., Adlof, S. M., Hogan, T. P., Ellis, A., and Weismer, S. (2005a). Are SLI and dyslexia distinct disorders? *J. Speech Lang. Hear Res.* 48, 1378–1396. doi: 10.1044/1092-4388(2005/096)
- Catts, H. W., Hogan, T. P., and Fey, M. (2003). Subgrouping poor readers on the basis of reading-related abilities. *J. Learn. Disabil.* 36, 151–164. doi: 10.1177/002221940303600208
- Catts, H. W., and Kahmi, A. (2005). *Language and Learning Disabilities*. Boston, MA: Pearson.
- Catts, H. W., Kahmi, A. G., and Adlof, S. M. (2005b). “Defining and classifying reading disabilities,” in *Language and Learning Disabilities*, eds H. W. Catts and A. G. Kahmi (Boston, MA: Pearson), 45–76.
- Coté, N., Goldman, S. R., and Saul, E. U. (1998). Students making sense of informational text: relations between processing and representation. *Discourse Process.* 25, 1–53. doi: 10.1080/01638539809545019
- Curran, C. E., Kintsch, E., and Hedberg, N. (1996). Learning-disabled adolescents’ comprehension of naturalistic narratives. *J. Educ. Psychol.* 88, 494–507. doi: 10.1037/0022-0663.88.3.494
- Denckla, M. B., and Rudel, R. (1974). Rapid “Automatized” naming of picture objects, colors, letters, and numbers by normal children. *Cortex* 10, 186–202. doi: 10.1016/S0010-9452(74)80009-2
- Ellis, A. W., McDougall, S., and Monk, A. F. (1996). Are dyslexics, different? A comparison between dyslexics, reading age controls, poor readers, and precocious readers. *Dyslexia Int. J. Pract. Res.* 2, 59–68. doi: 10.1002/(SICI)1099-0909(199602)2:1<31::AID-DYS34>3.3.CO;2-S
- Frazier, L., and Fodor, J. D. (1978). The sausage machine: a new two-stage parsing model. *Cognition* 6, 291–325. doi: 10.1016/0010-0277(78)90002-1
- Frith, U., and Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *Br. J. Educ. Psychol.* 1, 329–342.
- 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
- García Madruga, J. A., Martín Luque, L., and Santamaría, C. (1996). Intervención sobre la comprensión y recuerdo de textos: un programa de instrucción experimental. *Infanc. Aprendiz.* 74, 67–82. doi: 10.1174/021037096763000790
- Garnham, A., Oakhill, J., and Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. *Cognition* 11, 29–46. doi: 10.1016/0010-0277(82)90003-8
- Gonzales, A. J. C. A. (2008). *Lectura Conjunta, Pensamiento en Voz Alta y Comprensión Lectora*. doctoral tesis. Universidad de Salamanca, Salamanca.
- Hansen, C. L. (1978). Story retelling used with average and learning disabled readers as a measure of reading comprehension. *Learn. Disabil. Q.* 1, 62–69. doi: 10.2307/1510938
- Heaton, K. R., Chelune, G. J., Talley, J. L., Kay, G. G., and Curtiss, G. (2005). *Manual do Teste Wisconsin de Classificação de Cartas*. São Paulo: Casa do Psicólogo.
- Juel, C. (1988). Learning to read and write: a longitudinal study of 54 children from first to fourth grade. *J. Educ. Psychol.* 80, 437–447.
- Kessler, T. M. (1997). *Estudo da Memória Operacional em Pré-escolares*. Master tesis, Universidade Federal de Santa Maria, Santa Maria.
- Kida, A. S. B. (2015). *Sensibilidade e Especificidade da Tarefa de Reconto na Identificação de Marcadores Linguísticos e de Compreensão Leitora Para o Diagnóstico dos Transtornos da Leitura*. Relatório de Atividades: Conselho Nacional de Desenvolvimento Tecnológico (CNPQ).
- Kida, A. S. B., Bueno, G. J., Lima, V. L. C. C., Rossi, S. G., Nepomuceno, P. F., Martin, M. M. M., et al. (2015). Influência da modalidade de reconto na avaliação do desempenho de escolares em compreensão leitora. *Estudos Psicol.* 32, 605–615. doi: 10.1590/0103-166X2015000400004

- Kintsch, W. (1988). *Comprehension: A Paradigm for Cognition*. New York, NY: Cambridge University Press.
- Kintsch, W. (1998). The role of knowledge in discourse comprehension: a construction integration model. *Psychol. Rev.* 95, 163–182. doi: 10.1037/0033-295X.95.2.163
- Kintsch, W., and Keenan, J. M. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cogn. Psychol.* 5, 257–274. doi: 10.1016/0010-0285(73)90036-4
- Kornev, A. N., and Balciniene, I. (2014). *Story (re-)Telling and Reading in Children with Dyslexia: Language or Cognitive Resource Deficit?* *Academia*. Available at: [http://www.academia.edu/6796467/Story\\_re\\_telling\\_and\\_reading\\_in\\_children\\_with\\_dyslexia\\_language\\_or\\_cognitive\\_resource\\_deficit](http://www.academia.edu/6796467/Story_re_telling_and_reading_in_children_with_dyslexia_language_or_cognitive_resource_deficit)
- LaBerge, D., and Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. doi: 10.1016/0010-0285(74)90015-2
- León, J. A. (1991). La comprensión y recuerdo de textos expositivos a través del análisis de algunas variables del texto y lector. *Infanc. Aprendiz.* 14, 51–59. doi: 10.1080/02103702.1991.10822314
- Leslie, L., and Caldwell, J. (2009). “Formal and informal measures of reading comprehension,” in *Handbook of Research on Reading Comprehension*, eds S. E. Israel and G. G. Duffy (New York, NY: Taylor & Francis), 403–427.
- McNamara, D. S., Graesser, A. C., Cai, Z., and Kulikowich, J. M. (2012). “Coh-matrix easability components: aligning text difficulty with theories of text comprehension,” in *Proceedings of the Annual Meeting of the American Educational Research Association*, New Orleans.
- McNamara, D. S., Louwerse, M. M., McCarthy, P. M., and Graesser, A. C. (2010). Coh-matrix: capturing linguistic features of cohesion. *Discourse Processes*. 47, 292–330.
- Meyer, B. J. F. (1984). “Organizational aspects of text: effects on reading comprehension and applications for the classroom,” in *Promoting Reading Comprehension*, ed. J. Flood (Newark, DE: International Reading Association), 113–138.
- Miller-Shaul, S. (2005). The characteristics of young and adult dyslexics readers on reading and reading related cognitive tasks as compared to normal readers. *Dyslexia* 11, 132–151. doi: 10.1002/dys.290
- Nascimento, T. A., Carvalho, C. A. F., Kida, A. S. B., and Ávila, C. R. B. (2011). Fluência e compreensão leitora em escolares com dificuldades de leitura. *J. Soc. Bras. Fonoaudiol.* 23, 335–343.
- Nation, K., and Snowling, M. J. (1988). Individual differences in contextual facilitation: evidence from dyslexia and poor reading comprehension. *Child Dev.* 69, 996–1011. doi: 10.2307/1132359
- Oakhill, J. V., Yuill, N. M., and Parkin, A. J. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *J. Res. Read.* 9, 80–91. doi: 10.1111/j.1467-9817.1986.tb00115.x
- Orrantia, J., Sánchez, E. M., and Rosales, J. (1990). Hacia una medición de las estrategias implicadas en el proceso de comprensión lectora. *Lectura Vida*. 4, 24–31.
- Owens, J., Bower, G. H., and Black, J. B. (1979). The soap opera effect in story recall. *Mem. Cogn.* 7, 181–195. doi: 10.3758/BF03197537
- Perfetti, C. A. (1985). *Reading Ability*. New York, NY: Oxford University Press.
- Perfetti, C. A., and Lesgold, A. M. (1977). “Discourse comprehension and sources of individual differences,” in *Cognitive Processes in Comprehension*, eds M. A. Just and P. A. Carpenter (Hillsdale, NJ: Erlbaum).
- Perfetti, C. A., and Roth, S. (1981). “Some of the interactive processes in reading and their role in reading skill,” in *Interactive Processes in Reading*, eds A. Lesgold and C. Perfetti (Hillsdale, NJ: Erlbaum).
- Pinheiro, Â. M. V. (2011). “Transparência ortográfica e o efeito de retroalimentação fonológico grafêmica: implicações para a construção de provas de reconhecimento de palavras,” in *Dislexia: Novos temas, Novas Perspectivas*, eds L. M. Alves, R. Mousinho, and S. A. Capellini (Rio de Janeiro: Wak), 131–146.
- Puranik, C., Lombardino, L. J., and Altmann, L. J. (2006). Writing through retellings: an exploratory study of language-impaired and dyslexic populations. *Read Writ.* 20, 251–272. doi: 10.1007/s11145-006-9030-1
- Puranik, C. S., Lombardino, L. J., and Altmann, L. J. P. (2008). Assessing the microstructure of written language using a retelling paradigm. *Am. J. Speech Lang. Pathol.* 17, 107–120. doi: 10.1044/1058-0360(2008/012)
- Raven, J. C., Raven, J., and Court, J. H. (1988). *Matrizes Progressivas Coloridas de Raven - Manual*. São Paulo: Casa do Psicólogo.
- Reed, D. K., and Vaughn, S. (2012). Retell as an indicator of reading comprehension. *Sci. Stud. Read.* 16, 187–217. doi: 10.1080/1088438.2010.538780
- Richgels, D. J., McGee, L. M., Lomax, R. G., and Sheard, C. (1987). Awareness of text structures: effects on recall of expository text. *Read. Res. Q.* 22, 177–196. doi: 10.2307/747664
- Roller, C. M. (1990). The interaction between knowledge and structure variables in the processing of expository prose. *Read. Res. Q.* 25, 79–89. doi: 10.2307/747595
- Salgado, C. A., and Capellini, S. A. (2008). Programa de remediação fonológica em escolares com dislexia do desenvolvimento. *Pro Fono* 20, 31–36. doi: 10.1590/S0104-56872008000100006
- Sánchez Miguel, E. (2002). *Compreensão e Redação de Textos: Dificuldades e Ajudas*. Porto Alegre: Artmed.
- Scarton, C., and Aluísio, S. (2010). Análise da inteligibilidade de textos via ferramentas de processamento de língua natural: adaptando as métricas do Coh-matrix para o Português. *Linguamática* 2, 45–62.
- Shankweiler, D., Katz, L., Stuebing, K., Fletcher, J. M., Brady, S., Fowler, A., et al. (1999). Comprehension and decoding: patterns of association in children with reading difficulties. *Sci. Stud. Read.* 3, 69–94. doi: 10.1207/s1532799xssr0301\_4
- Shankweiler, D., and Liberman, I. Y. (1972). “Misreading: a search for causes,” in *Language by Ear and by Eye: The Relationships between Speech and Reading*, eds J. F. Kavanagh and G. Mattingly (Cambridge, MA: MIT Press).
- Siegel, L. S., and Ryan, E. B. (1988). Development of grammatical sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Dev. Psychol.* 20, 200–207. doi: 10.1037/0012-1649.20.2.200
- Simmons, F., and Singleton, C. (2000). The reading comprehension abilities of dyslexic students in higher education. *Dyslexia*. 6, 178–192. doi: 10.1002/1099-0909(200007/09)6:3<178::AID-DYS171>3.0.CO;2-9
- Simões, V. F. (2006). *Estudo do Desempenho de Crianças das Séries Iniciais do Ensino Fundamental I em Testes de Leitura, Escrita e Nomeação Rápida*. Master dissertation. Universidade de São Paulo, São Paulo.
- Snyder, L. S., and Downey, D. M. (1991). The language-reading relationship in normal and reading-disabled children. *J. Speech Lang. Hear. Res.* 34, 129–140. doi: 10.1044/jshr.3401.129
- Squires, K. E., Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., Bohman, T. M., and Gillam, R. B. (2014). Story retelling by bilingual children with language impairments and typically-developing controls. *Int. J. Lang. Commun. Disord.* 49, 60–74. doi: 10.1111/1460-6984
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the acquisition of literacy. *Read. Res. Quart.* 16, 32–71. doi: 10.2307/747348
- Stanovich, K. E., and Siegel, L. S. (1994). The phenotypic performance profile of reading-disabled children: 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
- Stein, L. M. (1994). *Teste de Desempenho Escolar: Manual Para Aplicação e Interpretação*. São Paulo: Casa do psicólogo.
- Tabaqui, M. L. M. (2008). *Validação do Exame Neuropsicológico e Análise das Funções Corticais Superiores em Crianças do Ensino Fundamental*. Doutorado. UCLA Unicamp, Campinas.
- Urbina, S. (2007). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artmed.
- Vogel, S. A. (1975). *Syntactic Abilities in Normal and Dyslexic Children*. Baltimore: University Park Press.
- Weaver, P. A., and Dickinson, D. K. (1979). Story comprehension and recall in dyslexic students. *Bull. Orton Soc.* 29, 157–171. doi: 10.1007/BF02653740
- Wechsler, D. (2002). “WISC III: wechsler intelligence scale for children-manual,” in *Adaptação e Padronização de Uma Amostra Brasileira*, 3rd Edn, ed. V. L. M. Figueiredo (São Paulo: Casa do Psicólogo).
- Williams, J. P. (1991). Comprehension by learning-disabled and nondisabled adolescents of personal/social problems presented in text. *Am. J. Psychol.* 104, 563–586. doi: 10.2307/1422941



Zinar, S. (1990). Fifth-graders' recall of propositional content and causal relationships from expository prose. *J. Read. Behav.* 22, 181–199.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2016 Kida, Ávila and Capellini. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# Orthographic Reading Deficits in Dyslexic Japanese Children: Examining the Transposed-Letter Effect in the Color-Word Stroop Paradigm

Shino Ogawa<sup>1\*</sup>, Masahiro Shibasaki<sup>2</sup>, Tomoko Isomura<sup>3</sup> and Nobuo Masataka<sup>4</sup>

<sup>1</sup> Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan, <sup>2</sup> Faculty of Intercultural Studies, Nagoya Gakuin University, Aichi, Japan, <sup>3</sup> Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan, <sup>4</sup> Section of Cognition and Learning, Primate Research Institute, Kyoto University, Aichi, Japan

## OPEN ACCESS

### Edited by:

Giseli Donadon Germano,  
Universidade Estadual Paulista, Brazil

### Reviewed by:

Angela Jocelyn Fawcett,  
Swansea University, UK  
Junhong Yu,  
The University of Hong Kong, China

### \*Correspondence:

Shino Ogawa  
ogawa.shino.57n@st.kyoto-u.ac.jp

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 19 November 2015

**Accepted:** 09 May 2016

**Published:** 31 May 2016

### Citation:

Ogawa S, Shibasaki M, Isomura T and  
Masataka N (2016) Orthographic  
Reading Deficits in Dyslexic Japanese  
Children: Examining the  
Transposed-Letter Effect in the  
Color-Word Stroop Paradigm.  
*Front. Psychol.* 7:767.  
doi: 10.3389/fpsyg.2016.00767

In orthographic reading, the transposed-letter effect (TLE) is the perception of a transposed-letter position word such as “chocolate” as the correct word “chocolate.” Although previous studies on dyslexic children using alphabetic languages have reported such orthographic reading deficits, the extent of orthographic reading impairment in dyslexic Japanese children has remained unknown. This study examined the TLE in dyslexic Japanese children using the color-word Stroop paradigm comprising congruent and incongruent Japanese hiragana words with correct and transposed-letter positions. We found that typically developed children exhibited Stroop effects in Japanese hiragana words with both correct and transposed-letter positions, thus indicating the presence of TLE. In contrast, dyslexic children indicated Stroop effects in correct letter positions in Japanese words but not in transposed, which indicated an absence of the TLE. These results suggest that dyslexic Japanese children, similar to dyslexic children using alphabetic languages, may also have a problem with orthographic reading.

**Keywords:** dyslexia, Japanese, orthographic reading, Stroop, transposed-letter effect

## INTRODUCTION

Dyslexia is a developmental disorder characterized by reading difficulty in children and adults of normal intelligence who have the motivation to read accurately and fluently (Shaywitz and Shaywitz, 2005). The prevalence of dyslexia varies depending on the linguistic system. For instance, dyslexia is estimated in ~5–12% of participants who used English as a primary language (Katusic et al., 2001). Dyslexia has also been found in participants with non-alphabetic languages, such as Japanese, but at much lower percentages (Uno et al., 2009). This suggests that differences in the architecture of English and Japanese may be associated with a propensity to dyslexia at least partly; therefore, research that compares dyslexia in different language systems may provide important insights into its mechanism.

The underlying mechanisms of dyslexia have thus far remained largely unclear (Gabrieli, 2009; Dehaene et al., 2010). According to Coltheart’s dual-route model (Coltheart et al., 2001), written words are processed in either lexical (orthographic) or sub-lexical (phonological) reading routes. For dyslexic users of alphabetic languages, both these routes are believed to be impaired

(Gabrieli, 2009; Peterson and Pennington, 2012). In phonological reading, dyslexia can cause deficits in both the segmentation of a speech stream into phonological units and the association of each unit with its corresponding letter (Shaywitz and Shaywitz, 2005). Japanese writing has three different character systems: hiragana, katakana, and kanji. Dyslexia has been estimated to occur in ~0.2, 1.4, and 6.9%, for each system respectively (Uno et al., 2009). Previous research has reported that Japanese school-age children with kana (hiragana and katakana) dyslexia have difficulty associating each phonological unit with its corresponding letter but have no difficulty segmenting the speech stream into phonological units (Ogawa et al., 2014). This suggests that different mechanisms may be involved in English and Japanese dyslexia. These could be associated with some of the characteristics of kana, such as its psycholinguistic grain size and orthography-to-phonology translation relationships, that are quite different from alphabets (Wydell and Butterworth, 1999).

Previous studies involving alphabetic languages have reported orthographic process deficits in dyslexics (O'Brien et al., 2011; Kezilas et al., 2014; Ziegler et al., 2014). Some Japanese dyslexics have also exhibited difficulty in reading texts after acquiring a reading knowledge of kana characters (Yoshida and Tsuzuki, 2015). As most kana characters express one sound (Wydell and Butterworth, 1999), any difficulties faced after acquiring a reading knowledge of kana characters cannot be explained by phonological reading deficits. One possibility is that dyslexic children may also face difficulties with the orthographic process; that is, recognizing a word as a whole, along the lexical reading route. However, whether orthographic reading is also impaired in Japanese dyslexics remains unclear. In orthographic reading, letter positions within a word can be loosely perceived (Carreiras et al., 2015). For instance, a transposed-letter nonword such as “cholocate” is frequently misperceived as the word “chocolate.” This transposed-letter effect (TLE) has been reported in various European and non-European alphabetic languages, such as English (Perea and Lupker, 2003; Johnson et al., 2007), French (Schoonbaert and Grainger, 2004), Spanish (Perea and Carreiras, 2006a,b), and Basque (Duñabeitia et al., 2007). The TLE can be assessed using the color-word Stroop paradigm (Arsalidou et al., 2013). Arsalidou et al. (2013) revealed that the Stroop effect, that is, the interference observed when a color word and the actual printed color of the word are incongruent (e.g., the word “red” printed in blue), was observed in correct words (e.g., purple) as well as in transposed-letter nonwords (e.g., prulpe) in English (Arsalidou et al., 2013). Therefore, orthographic reading deficits can be evaluated by the presence or absence of the TLE in the color-word Stroop paradigm.

In this study, we examined orthographic reading deficits in dyslexic Japanese children with TLE using the color-word Stroop paradigm. To confirm the suitability of this experimental procedure (Experiment 1), we first examined whether TLE was observed in normal Japanese adults reading Japanese kana words using the color-word Stroop test. Thereafter, we examined orthographic reading impairments in dyslexic Japanese children (Experiment 2). If these children suffered from orthographic reading deficits, they were not expected to display or display a small TLE compared with normal Japanese children.

## EXPERIMENT 1

### Materials and Methods

#### Ethics Note

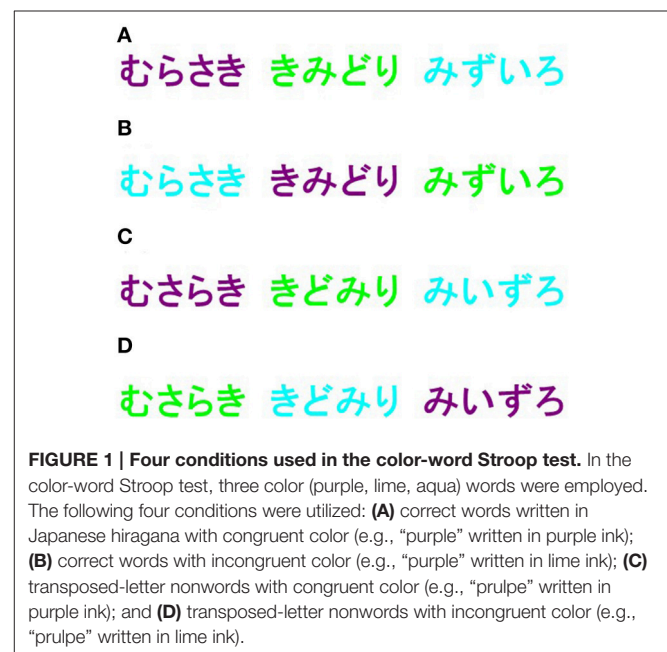
This study was conducted in accordance with the principles expressed in the Declaration of Helsinki and the Ethical Guidelines for Medical and Health Research Involving Human Subjects by the Japanese Ministry of Health, Labour, and Welfare. All experimental protocols were approved by the Institutional Ethics Committee of the Primate Research Institute, Kyoto University (permission number, H2012-09).

#### Participants

Participants included 22 Japanese adults (11 males and 11 females; mean age = 26.37; *SD* = 3.55) with no psychiatric or neurological conditions. All had normal or corrected-to-normal visual acuity and adequate color vision. All participants provided informed written consent to participate in this study.

#### Color-Word Stroop Test

In the color-word Stroop test (Stroop, 1935), participants are asked to name the ink color in which a congruent color word is written (i.e., the word “red” written in red ink: congruent condition) or the ink color of an incongruent color word (i.e., the word “blue” written in red ink: incongruent condition). Stroop effects are determined by comparing the reaction time (RT) in the incongruent condition with the RT in the congruent condition. Three colors were used for the Japanese characters in this test. Given that creating transposed-letter nonwords requires a relatively longer word length (four Japanese characters), purple [RGB (128, 0, 128)], lime [RGB (0, 255, 0)], and aqua [RGB (0, 255, 255)] were selected, which are commonly used colors that all participants were familiar with (Figure 1).



All participants could read and name the colors accurately. The test was conducted as a computerized manual response paradigm controlled by a custom-written software, Visual Basic 6.0 (Microsoft Corporation, Redmond, Washington, USA), and was run on a personal computer so that each participant pressed a button in response when they saw the stimulus on the screen. In this study, the RT was measured using a manual response paradigm.

To familiarize the participants with the test, one 20-trial training session was conducted before the actual test session. Stimuli created with MS P Gothic font, size 72, bold style, were presented without time limits in the center of a color monitor's visual field ~30 cm from the participants. During inter-trial intervals, a fixation mark (a white plus sign) was shown in the center of the visual field. Inter-trial intervals varied pseudo-randomly between 1500 and 2000 ms (increments of 100 ms). Responses were provided by pressing pre-defined buttons on the 10-digit keyboard without feedback. Colored stickers were placed on the relevant keys to reduce memory demands. Each participant's matching of colored stickers and key positions were randomly assigned. Participants were instructed to respond to the ink color of the stimuli as quickly and as accurately as possible. Stimuli included the following four conditions: (a) correct words written in Japanese hiragana with congruent color (e.g., "purple" written in purple ink); (b) correct words with incongruent color (e.g., "purple" written in lime ink); (c) transposed-letter nonwords with congruent color (e.g., "prulpe" written in purple ink); and (d) transposed-letter nonwords with incongruent color (e.g., "prulpe" written in lime ink) (**Figure 1**). In the transposed-letter nonwords, the word's first and last letters were kept in place while the middle letters were transposed. The four different stimuli conditions were pseudo-randomly presented with equal frequency (18 trials for each condition), thus resulting in 72 trials per session. Both accuracy and RTs were recorded.

## Data Analysis

Mean RTs were calculated for the trials with correct responses. RTs in trials greater and smaller than 2 SD from each participant's

mean were eliminated from subsequent analysis as outliers (5.05%,  $SD = 1.39$ ).

The Stroop effect, which was used as a dependent variable in the statistical analyses, was calculated using the following formula:  $[(\text{Incongruent RT} - \text{Congruent RT}) / \text{Incongruent RT} \times 100]$  (Mayas et al., 2012),  $[(\text{Incongruent RT} - \text{Congruent RT}) / \text{Congruent RT} \times 100]$  (Naccache et al., 2005).

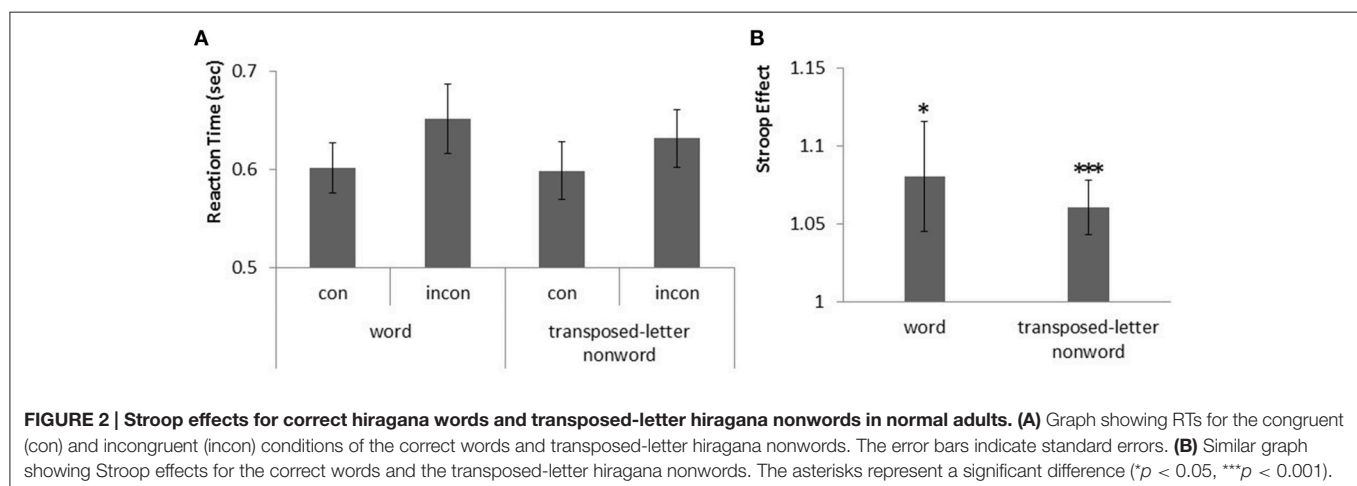
The manipulation of the letter positions that affected the Stroop effect were examined using an analysis of covariance (ANCOVA) on the Stroop effects for the correct and transposed-letter words, with the WORD CONDITION (correct word vs. transposed-letter nonword) as the fixed within-subject factor and controlled for gender and individual error rate. Furthermore, to examine whether the Stroop effect itself existed, we ran a one-sample  $t$ -test on the Stroop effect for the correct or transposed-letter words. Statistical analyses were conducted using the freeware "R 3.2.3" (R Development Core Team) as "SPSS Statistics 22" (IBM Japan, Ltd.).

## Results

The average error rate was 3.47% ( $SD = 4.28$ ). The mean RTs with correct and transposed-letter positions are shown in **Figure 2A**, and the Stroop effects for correct words and transposed-letter nonwords are shown in **Figure 2B**.

The ANCOVA results were as follows: as a parameter, gender showed no significance in the regression for both correct word condition [ $\beta = 0.08$ ,  $p = 0.24$ ] and transposed-letter nonword condition [ $\beta = 0.04$ ,  $p = 0.21$ ]; the error rate also showed no significance in the regression for both correct word condition [ $\beta = 1.34$ ,  $p = 0.11$ ] and transposed-letter nonword condition [ $\beta = 0.36$ ,  $p = 0.40$ ].

As no significance was observed in the regression for either gender error rate, we ran a paired  $t$ -test on the Stroop effects of the correct and transposed-letter words with WORD CONDITION (correct word vs. transposed-letter nonword). The Stroop effects showed no difference between the word and transposed-letter nonword [ $t_{(21)} = 0.61$ ,  $p = 0.54$ ].





The one-sample *t*-test revealed statistically significant Stroop effects in both the correct words [ $t_{(21)} = 2.27, p = 0.01$ , Cohen's  $d = 0.34$ ] and the transposed-letter words [ $t_{(21)} = 3.43, p = 0.001$ , Cohen's  $d = 0.23$ ].

## Discussion

In this experiment, Stroop effects were observed for both correct words and transposed-letter nonwords, and neither error rate nor gender confounded the Stroop effects in adults. Therefore, these results suggested that normal Japanese adults utilize orthographic reading when recognizing transposed-letter nonwords in the color-word Stroop paradigm.

## EXPERIMENT 2

### Materials and Methods

#### Participants

Participants were 20 typically developing (TD) children (11 male and 9 female) (mean age = 11.27;  $SD = 1.01$ ; range = 10.08–13.00) without psychiatric or neurological conditions, and 11 dyslexic children (7 male and 4 female) (mean age = 11.65;  $SD = 0.35$ ; range = 10.91–13.00). **Table 1** summarizes the information on the dyslexic children, indicating their gender, full-scale IQ, and age. All participants had been diagnosed by a child psychiatrist at a general hospital or a child consultation center, and all demonstrated difficulty in reading Japanese compared to their peers. They had been receiving special training because of their dyslexia for 1 to 4 years, and they could therefore read hiragana words well when there were limited words shown; however, they could not correctly read many words in longer texts. The Intelligence Quotient (IQ) was measured using the Japanese version of the Wechsler Intelligence Scale for Children (either WISC-III or WISC-IV). Age [ $t_{(29)} = -1.14, p = 0.26$ ] and Full-scale IQ [ $t_{(29)} = 1.94, p = 0.06$ ] and was the same for both the TD and dyslexic participants. All participants had normal or corrected-to-normal visual acuity and adequate color vision. The participants' parents provided informed written consent for their children's participation in this study.

**TABLE 1 | Information on the dyslexic children in the study.**

ID	Gender	FIQ	Age
1	Female	101	13 y 0 m
2	Male	100	12 y 0 m
3	Male	104	12 y 0 m
4	Male	79	12 y 0 m
5	Male	103	11 y 11 m
6	Female	90	11 y 6 m
7	Male	97	11 y 4 m
8	Female	101	11 y 3 m
9	Female	112	11 y 2 m
10	Male	88	11 y 2 m
11	Male	93	10 y 11 m

### Color-Word Stroop Test

The same color-word Stroop test as described in Experiment 1 was used. All participants could accurately read and name the colors. When the children demanded breaks during the test, short breaks of no longer than 10 min were given.

### Data Analysis

Mean RTs were calculated for trials with correct responses. RTs  $> 4$  s and RTs in trials greater and smaller than 2 SD from each participant's mean were eliminated from subsequent analysis as outliers (TD children: 4.72%,  $SD = 1.71$ ; dyslexic children: 7.07%,  $SD = 2.73$ ).

The Stroop effects as a dependent variable were calculated using the following formula: [Incongruent RT/Congruent RT] (cf. [(Incongruent RT – Congruent RT)/Incongruent RT  $\times$  100] (Mayas et al., 2012), [(Incongruent RT – Congruent RT)/Congruent RT  $\times$  100] (Naccache et al., 2005).

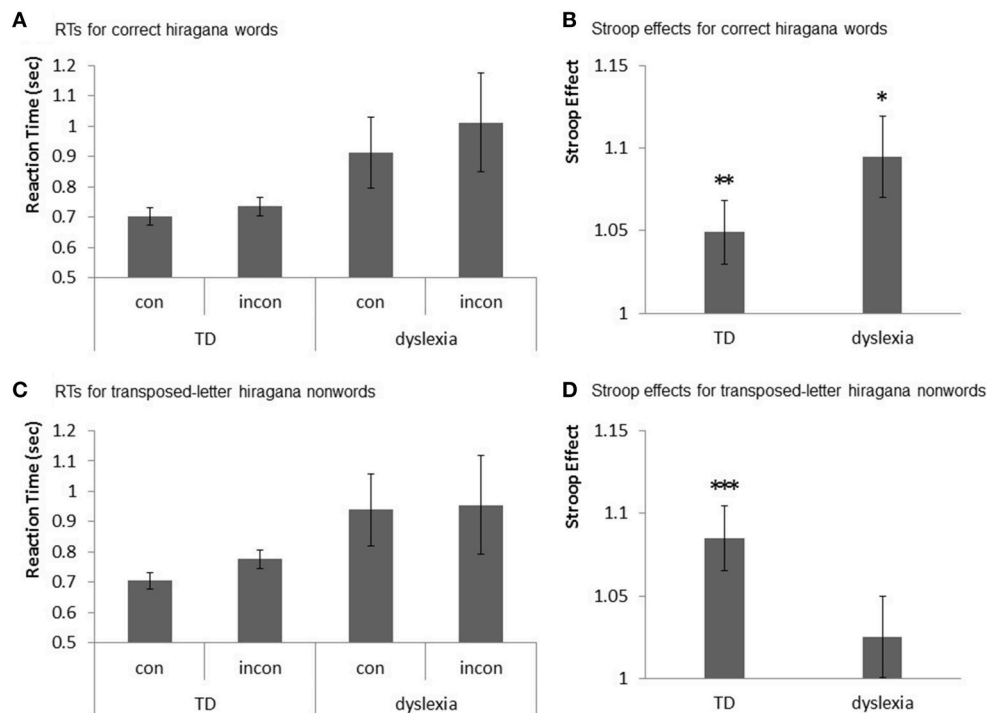
To examine the Stroop effect in each word condition for the TD/dyslexic participants, we ran an ANCOVA on the Stroop effects of the correct and transposed-letter words with WORD CONDITION (correct word vs. transposed-letter nonword) as the fixed within-subject factor and GROUP CONDITION (TD vs. dyslexia) as the fixed between-subject factor, with gender, error rate, FIQ, and age as covariates. To examine whether the Stroop effect existed, we ran a one-sample *t*-test on the Stroop effect of the correct or transposed-letter words. Statistical analyses were conducted using the freeware “R 3.2.3” (R Development Core Team) as “SPSS Statistics 22” (IBM Japan, Ltd).

## Results

The average error rates were 4.44% ( $SD = 3.32$ ) in the TD children and 4.29% ( $SD = 5.57$ ) in the dyslexic children; therefore, no significant difference was observed between the groups [ $t_{(29)} = -0.09, p = 0.92$ ].

Mean correct word RTs for the TD and dyslexic children are shown in **Figure 3A**, and the Stroop effects in the TD and dyslexic children are shown in **Figure 3B**. In addition, Mean transposed-letter nonwords RTs for the TD and dyslexic children are shown in **Figure 3C**, and the Stroop effects in TD and dyslexic children are shown in **Figure 3D**.

The ANCOVA results were as follows. There were regression parallels between GROUP CONDITION and gender [ $F_{(1, 21)} = 2.50, p = 0.12$ ], GROUP CONDITION and error rate [ $F_{(1, 21)} = 0.40, p = 0.53$ ], GROUP CONDITION, and FIQ [ $F_{(1, 21)} = 0.29, p = 0.59$ ], and GROUP CONDITION and age [ $F_{(1, 21)} = 0.28, p = 0.59$ ]. Gender showed no significance in the regression for the transposed-letter nonword condition [ $\beta = -0.015, p = 0.70$ ]; however, it was significant in the regression for the correct word condition [ $\beta = -0.068, p = 0.09$ ]. Error rate, similarly, showed no significance in the regression for the transposed-letter nonword condition [ $\beta = 0.134, p = 0.24$ ] but was significant in the regression for the correct word condition [ $\beta = -0.794, p = 0.07$ ]. The FIQ showed no significance in the regression for both the correct word condition [ $\beta = 0.002, p = 0.27$ ] and the transposed-letter nonword condition [ $\beta = 0.000, p = 0.83$ ]. Age showed no significance in the regression for the correct



**FIGURE 3 | Stroop effects for correct hiragana words and transposed-letter hiragana nonwords in TD and dyslexic children. (A)** Graph showing the RTs for the congruent (con) and incongruent (incon) conditions of the correct hiragana words with TD and dyslexic children. **(B)** Graph showing the Stroop effects for the correct hiragana words with TD and dyslexic children. The asterisks represent a significant difference (\* $p < 0.05$ , \*\* $p < 0.01$ ). **(C)** Graph showing the RTs for the congruent (con) and incongruent (incon) conditions for the transposed-letter hiragana nonwords with TD and dyslexic children. **(D)** Graph showing the Stroop effects for the transposed-letter hiragana nonwords in TD children, but not in dyslexic children. The asterisk represents a significant difference (\*\* $p < 0.001$ ).

word condition [ $\beta = -0.003$ ,  $p = 0.87$ ] but was significant in the regression for the transposed-letter nonword condition [ $\beta = -0.044$ ,  $p = 0.04$ ].

As no significance was observed in the FIQ in the regression, the interaction between WORD CONDITION and gender, error rate, or age, we ran an ANOVA on the Stroop effects of the correct and transposed-letter words with WORD CONDITION (correct word vs. transposed-letter nonword) as the within-subject factor and GROUP CONDITION (TD vs. dyslexia) as the between-subject factor.

**Table 2** summarizes the variance sources. No main effect was observed for GROUP CONDITION [ $F_{(1, 29)} = 0.36$ ,  $p = 0.55$ ,  $\eta^2_G = 0.004$ ] or WORD CONDITION [ $F_{(1, 29)} = 0.06$ ,  $p = 0.80$ ,  $\eta^2_G = 0.001$ ]. A significant trend was observed in the interaction between GROUP CONDITION and WORD CONDITION [ $F_{(1, 29)} = 3.72$ ,  $p = 0.06$ ,  $\eta^2_G = 0.039$ ].

The simple main effect of the WORD CONDITION was not significant for either the TD children [ $F_{(1, 29)} = 1.97$ ,  $p = 0.17$ ] or the dyslexic children [ $F_{(1, 29)} = 1.84$ ,  $p = 0.18$ ]. However, the simple main effect of the GROUP CONDITION was marginally significant for the transposed-letter nonword condition [ $F_{(1, 29)} = 3.43$ ,  $p = 0.07$ ] but not for the correct word condition [ $F_{(1, 29)} = 0.77$ ,  $p = 0.38$ ].

We used a one-sample  $t$ -test to examine whether the Stroop effect existed for each condition. For the correct word, the one-sample  $t$ -test revealed statistically significant effects in both the TD children [ $t_{(19)} = 2.50$ ,  $p = 0.01$ , Cohen's  $d = 0.24$ ] and the

**TABLE 2 | Summary of variance sources tested using a TD and dyslexic children ANOVA.**

Source	df	F	p	$\eta^2_G$
GROUP CONDITION	1	0.36	0.55	0.004
WORD CONDITION	1	0.06	0.80	0.001
GROUP CONDITION $\times$ WORD CONDITION	1	3.72	0.06	0.039

The symbol  $\eta^2_G$  represents generalized eta squared statistics.

dyslexic children [ $t_{(10)} = 2.02$ ,  $p = 0.03$ , Cohen's  $d = 0.20$ ]. For the transposed-letter nonword, the one-sample  $t$ -test revealed statistically significant effects in the TD children [ $t_{(19)} = 3.87$ ,  $p = 0.0005$ , Cohen's  $d = 0.46$ ] but not in the dyslexic children ( $t_{(10)} = 1.06$ ,  $p = 0.15$ , Cohen's  $d = 0.03$ ).

## Discussion

In Experiment 2, TD and dyslexic children showed different Stroop effect tendencies for the correct word and transposed-letter nonword tasks. Dyslexic children showed a Stroop effect in the correct word condition but not in the transposed-letter nonword condition; however, the TD children indicated clear Stroop effects for both the word and transposed-letter nonword conditions. This result suggests that the dyslexic children who participated in this study can use phonological reading but face problems with orthographic reading. This does not contradict the fact that the dyslexic children participating in this study could

read hiragana words well when there were limited words shown but could not read several words in longer texts correctly.

If the hypothesis that dyslexic children were not expected to display or display a small TLE compared with normal Japanese children is true, the interaction between GROUP CONDITION and WORD CONDITION should be significant. In this experiment, the interaction was marginally significant but failed to achieve statistical significance. This may be attributed to the small sample size or large variance of dyslexic children. Because of many individual differences in dyslexic children, larger sample sizes will be required in future studies.

## GENERAL DISCUSSION

In this study, we found that both normal adults and TD children exhibited Stroop effects in Japanese hiragana words with transposed-letter positions, indicating the presence of TLE in orthographic reading. In contrast, Stroop effects for words with transposed-letter positions were not observed in dyslexic children, which was consistent with previous research on dyslexic children in other languages (O'Brien et al., 2011; Kezilas et al., 2014; Ziegler et al., 2014) and indicated that dyslexic Japanese children may also face difficulties in orthographic reading.

The current study demonstrated the TLE in Japanese hiragana using the Stroop paradigm, as has been previously shown for English (Arsalidou et al., 2013). Although, the unit size (syllable) of hiragana is greater than the unit size (phoneme) for the Roman alphabet, the letter position of a hiragana word may affect its orthographic readability similarly to that of Western languages that use the Roman alphabet. Consistent with our findings, two previous studies have examined TLE in Japanese kana using different experimental paradigms (Perea and Perez, 2009; Perea et al., 2011). Using a masked priming lexical decision paradigm, Perea and Perez (2009) reported that the lexical decision time for the word “a.me.ri.ka [アメリカ]” was faster when the prime was a transposed-letter nonword “a.ri.me.ka [アリメカ]” than the control nonword “a.ka.ho.ka [アカホカ].” Perea et al. (2011) also found that while using the silent reading paradigm, fixation time on the target word was shorter when the parafoveal preview was the transposed-letter nonword (a.ri.me.ka [アリメカ]–a.me.ri.ka [アメリカ]) than the control nonword (a.ka.ho.ka [アカホカ]–a.me.ri.ka [アメリカ]). A major limitation of these studies was that the experimental procedures often used unfamiliar words, especially for young children, it was therefore difficult to distinguish whether the low task performance was due to orthographic reading deficits or whether the words used in the task were unfamiliar to the participants. The color-word Stroop paradigm overcomes this problem as the words used in this paradigm can be fixed to only include familiar words.

People showing TLE can recognize letters with a high abstraction level. For people who have difficulty with the orthographic process, it may be difficult to recognize the same

character or letters in different fonts or handwritten characters. This could be supporting evidence for why dyslexics have difficulty reading even after they have acquired the characters or words.

This is the first report suggesting that orthographic reading is possibly impaired in Japanese dyslexics. From this and our previous study (Ogawa et al., 2014), Japanese dyslexics have been found to struggle with both phonological and orthographic reading, as in dyslexics from alphabetic language backgrounds (Gabrieli, 2009; Peterson and Pennington, 2012), although the percentage of dyslexics in Japan is much lower than in countries with speakers of alphabetic languages (Uno et al., 2009).

In particular, these results are very important for understanding and supporting dyslexic Japanese children who have difficulty reading texts even after having learnt how to read the kana characters, which cannot be explained by phonological reading deficits. This study suggests that these children may have difficulty with the orthographic process and may need special support specific to whole word recognition. This study included a small sample of dyslexic children. In the future, to ensure generalizability, larger sample sizes, and additional tests are required.

In conclusion, our study suggested that dyslexic Japanese children have difficulty in orthographic reading and that the Stroop paradigm was a useful tool in assessing the orthographic process for Japanese dyslexics.

## AUTHOR CONTRIBUTIONS

SO designed the study, and prepared the manuscript. MS programmed the computer-based task. SO, MS, and TI conducted the study. SO, MS, TI, and NM analyzed the data. SO, MS, TI, and NM prepared the manuscript.

## ACKNOWLEDGMENTS

This work was supported by the Grant-in-Aid for JSPS Fellows (11J05191 and 14J00317 to SO, 11J05358 to MS, and 12J03878 to TI). The participants in this research were recruited in a follow-up support program for children who had joined the learning support program by Nagoya City Child Welfare Center and the project on Developmental Disorders and Support for Acquiring Learning Skills and Communication by the Kokoro Research Center, Kyoto University. We would like to thank Dr. Yukiori Goto for his valuable comments and suggestions in data analysis and manuscript improvement, Ms. Satoko Yamada, Ms. Satomi Araya, Dr. Miwa Fukushima-Murata, Dr. Namiko Kubo-Kawai, Dr. Tomoko Asai, Dr. Hiroko Taniai, Dr. Yasuko Funabiki, Dr. Sakiko Yoshikawa, and members of the Nagoya City Child Welfare Center and Kokoro Research Center for their enormous help with the study and the participants for their cooperation. We would also like to thank Enago ([www.enago.jp](http://www.enago.jp)) for its English language review. Finally, we would like to express our sincere gratitude to the participants.

## REFERENCES

- Arsalidou, M., Agostino, A., Maxwell, S., and Taylor, M. J. (2013). I can read these colors. Orthographic manipulations and the development of the color-word stroop. *Front. Psychol.* 3:594 doi: 10.3389/fpsyg.2012.00594
- Carreiras, M., Quiñones, I., Hernández-Cabrera, J. A., and Duñabeitia, J. A. (2015). Orthographic coding: brain activation for letters, symbols, and digits. *Cereb. Cortex* 25, 4748–4760. doi: 10.1093/cercor/bhu163
- 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
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., NunesFilho, G., Jobert, A., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364. doi: 10.1126/science.1194140
- Duñabeitia, J. A., Perea, M., and Carreiras, M. (2007). Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. *Cognition* 105, 691–703. doi: 10.1016/j.cognition.2006.12.001
- Gabrieli, J. D. (2009). Dyslexia: a new synergy between education and cognitive neuroscience. *Science* 325, 280–283. doi: 10.1126/science.1171999
- Johnson, R. L., Perea, M., and Rayner, K. (2007). Transposed-letter effects in reading: evidence from eye movements and parafoveal preview. *J. Exp. Psychol. Hum.* 33, 209–229. doi: 10.1037/0096-1523.33.1.209
- Katusic, S. K., Colligan, R. C., Barbaresi, W. J., Schaid, D. J., and Jacobsen, S. J. (2001). Incidence of reading disability in a population-based birth cohort, 1976–1982, Rochester, Minn. *Mayo Clin. Proc.* 76, 1081–1092. doi: 10.4065/76.11.1081
- 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
- Mayas, J., Fuentes, L. J., and Ballesteros, S. (2012). Stroop interference and negative priming (NP) suppression in normal aging. *Arch. Gerontol. Geriatr.* 54, 333–338. doi: 10.1016/j.archger.2010.12.012
- Naccache, L., Dehaene, S., Cohen, L., Habert, M. O., Guichart-Gomez, E., Galanaud, D., et al. (2005). Effortless control: executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia* 43, 1318–1328. doi: 10.1016/j.neuropsychologia.2004.11.024
- O'Brien, B. A., Wolf, M., Miller, L. T., Lovett, M. W., and Morris, R. (2011). Orthographic processing efficiency in developmental dyslexia: an investigation of age and treatment factors at the sublexical level. *Ann. Dyslexia* 61, 111–135. doi: 10.1007/s11881-010-0050-9
- Ogawa, S., Fukushima-Murata, M., Kubo-Kawai, N., Asai, T., Taniai, H., and Masataka, N. (2014). Characteristics of developmental dyslexia in Japanese Kana: from the viewpoint of the Japanese feature. *J. Psychol. Abnorm. Child* 3:126. doi: 10.4172/2329-9525.1000126
- Perea, M., and Carreiras, M. (2006a). Do transposed-letter effects occur across lexeme boundaries? *Psychon. B. Rev.* 13, 418–422. doi: 10.3758/BF03193863
- Perea, M., and Carreiras, M. (2006b). Do transposed-letter similarity effects occur at a syllable level? *Exp. Psychol.* 53, 308–315. doi: 10.1027/1618-3169.53.4.308
- Perea, M., and Lupker, S. J. (2003). “Transposed-letter confusability effects in masked form priming,” in *The Masked priming: State of the Art*, ed S. Kinoshita and S. J. Lupker (New York, NY: Psychology Press), 97–120.
- Perea, M., Nakatani, C., and van Leeuwen, C. (2011). Transposition effects in reading Japanese Kana: are they orthographic in nature? *Mem. Cognit.* 39, 700–707. doi: 10.3758/s13421-010-0052-1
- Perea, M., and Perez, E. (2009). Beyond alphabetic orthographies: the role of form and phonology in transposition effects in Katakana. *Lang. Cognit. Proc.* 24, 67–88. doi: 10.1080/01690960802053924
- Peterson, R. L., and Pennington, B. F. (2012). Developmental dyslexia. *Lancet.* 379, 1997–2007. doi: 10.1016/S0140-6736(12)60198-6
- Schoonbaert, S., and Grainger, J. (2004). Letter position coding in printed word perception: effects of repeated and transposed letters. *Lang. Cognit. Proc.* 19, 333–367. doi: 10.1080/01690960344000198
- Shaywitz, S. E., and Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biol Psychiatry* 57, 1301–1309. doi: 10.1016/j.biopsych.2005.01.043
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Uno, A., Wydell, T. N., Haruhara, N., Kaneko, M., and Shinya, N. (2009). Relationship between reading/writing skills and cognitive abilities among Japanese primary-school children: normal readers versus poor readers (dyslexics). *Read. Writ.* 22, 755–789. doi: 10.1007/s11145-008-9128-8
- Wydell, T. N., and Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition* 70, 273–305. doi: 10.1016/S0010-0277(99)00016-5
- Yoshida, Y., and Tsuzuki, S. (2015). A study on diagnosis for the child with dyslexia [written in Japanese]. *Bull. Spec. Educ. Welf.* 11, 89–95. Available online at: <http://hdl.handle.net/10424/6133>
- Ziegler, J. C., Perry, C., and Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20120397. doi: 10.1098/rstb.2012.0397

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Ogawa, Shibasaki, Isomura and Masataka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# On the Development of Parafoveal Preprocessing: Evidence from the Incremental Boundary Paradigm

Christina Marx, Florian Hutzler\*, Sarah Schuster and Stefan Hawelka

Centre for Cognitive Neuroscience, University of Salzburg, Salzburg, Austria

## OPEN ACCESS

### Edited by:

Simone Aparecida Capellini,  
São Paulo State University "Júlio  
de Mesquita Filho", Brazil

### Reviewed by:

Thomas James Lundy,  
Virtuallaboratory.Net, Inc., USA

Rosa K. W. Kwok,  
University of Plymouth, UK

### \*Correspondence:

Florian Hutzler  
florian.hutzler@sbg.ac.at

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 30 November 2015

**Accepted:** 29 March 2016

**Published:** 14 April 2016

### Citation:

Marx C, Hutzler F, Schuster S  
and Hawelka S (2016) On  
the Development of Parafoveal  
Preprocessing: Evidence from  
the Incremental Boundary Paradigm.  
Front. Psychol. 7:514.  
doi: 10.3389/fpsyg.2016.00514

Parafoveal preprocessing of upcoming words and the resultant preview benefit are key aspects of fluent reading. Evidence regarding the development of parafoveal preprocessing during reading acquisition, however, is scarce. The present developmental (cross-sectional) eye tracking study estimated the magnitude of parafoveal preprocessing of beginning readers with a novel variant of the classical boundary paradigm. Additionally, we assessed the association of parafoveal preprocessing with several reading-related psychometric measures. The participants were children learning to read the regular German orthography with about 1, 3, and 5 years of formal reading instruction (Grade 2, 4, and 6, respectively). We found evidence of parafoveal preprocessing in each Grade. However, an effective use of parafoveal information was related to the individual reading fluency of the participants (i.e., the reading rate expressed as words-per-minute) which substantially overlapped between the Grades. The size of the preview benefit was furthermore associated with the children's performance in rapid naming tasks and with their performance in a pseudoword reading task. The latter task assessed the children's efficiency in phonological decoding and our findings show that the best decoders exhibited the largest preview benefit.

**Keywords:** reading fluency, reading acquisition, eye movement control during reading, incremental boundary paradigm, visual word recognition

## INTRODUCTION

While our eyes move across continuous texts in a sequence of fixations, we extract information not only from the word which we are currently fixating, but also from the not-yet fixated, upcoming word (Rayner, 1998). This parafoveal preview gives us first orthographic and phonological (and potentially lexical) information about the upcoming word (Schotter et al., 2012). Parafoveal preprocessing therefore accelerates foveal word recognition and hence contributes to fluent reading. Evidence regarding the developmental trajectory of parafoveal preprocessing, however, is limited.

Two gaze-contingent techniques are commonly used for investigating parafoveal preprocessing: (i) the *moving window paradigm* (McConkie and Rayner, 1975) and (ii) the *invisible boundary paradigm* (Rayner, 1975). Within the moving window paradigm, a text outside a predefined "window" to the left and right of the current fixation is masked, for example, by Xs. The text within the window is presented unmutated. By means of this paradigm, a reader's perceptual span can be estimated, that is, the minimal window size by which the reader is not affected by the parafoveal masks. Research using this paradigm demonstrated that the perceptual span for adult readers ranges from 3 to 4 letters left and 14 to 15 letters right of fixation

(e.g., McConkie and Rayner, 1975). By contrast, the perceptual span of beginning readers undergoes development, that is, it increases with reading experience. To illustrate, 2nd and 4th Grade children have a smaller span compared to adults, that is, about 3–4 letters to the left and about 11 letters to the right of fixation. Children from Grade 6, however, already show an adult-like span size (Rayner, 1986; Häikiö et al., 2009; Sperlich et al., 2015). In sum, evidence from the moving window paradigm suggests that children utilize information beyond the currently fixated word.

The most commonly used technique to study effects of parafoveal preprocessing of the upcoming word is the invisible boundary paradigm (Rayner, 1975). Within this paradigm an invisible boundary is placed before a theoretically relevant target word. As long as the reader fixates to the left of the boundary, a valid or an experimentally manipulated preview is presented (e.g., a X-mask, that is, a string of X's preserving the length of the target word or a same-shape/different-letter mask, that is, a sequence of different letters preserving the target word's length and shape). Contingent on crossing the boundary, the manipulated parafoveal preview is replaced with the target word. In order to estimate the preview benefit, fixation durations for valid previews are compared to those of manipulated (e.g., X-masked) previews. Research utilizing this paradigm showed – for adult, proficient readers – that the magnitude of the preview benefit is around 30–50 ms (Rayner, 2009).

Recent findings, however, indicated that the classical variant of the boundary paradigm does not provide an accurate estimate of the preview benefit (Hutzler et al., 2013; Kliegl et al., 2013; Marx et al., 2015). To be specific, when parafoveal masks are used as a baseline, they inflict processing costs and hence inflate the estimated preview benefit. A recent study from our lab revealed such an erroneous overestimation of the preview benefit in beginning readers (Marx et al., 2015). In the light of these recent findings, we adapted the classical approach and introduced the *incremental boundary technique* for investigating the development of parafoveal preprocessing in children (Marx et al., 2015). In short, instead of using parafoveal masks, we manipulated the salience of the parafoveal previews by gradually reducing its visual integrity (i.e., displacing a certain amount of pixels of the preview). In so doing, we can assess whether increasing salience leads to shorter processing times, that is, to a preview benefit (see Jacobs et al., 1995 for the logic of this within-condition baseline).

To date, three studies, which used the classical variant of the invisible boundary paradigm, provided evidence on parafoveal preprocessing in children. One study examined whether children from Grade 2, 4, and 6 extract information from a second constituent of a compound word (e.g., *ball* in *basketball*; the boundary was between *basket* and *ball*; Häikiö et al., 2010). This condition was compared to a condition which presented (space-separated) adjective–noun pairs (e.g., *little ball*). The authors reported that even 2nd Graders profited (in terms of shorter subsequent fixations) from “parafoveal” information when it was connected to the fixated word (i.e., the compound condition) compared to the adjective–noun condition. Another study examined whether 8 to 9-year-old children benefit

from parafoveal phonological information (i.e., by presenting pseudohomophone previews) and orthographic information (i.e., by presenting transposed-letter previews; Tiffin-Richards and Schroeder, 2015). They found that children – in contrast to adults – showed a pseudohomophone preview benefit, that is, they profited from the availability of phonological information in the parafovea. The third study investigated – in 4th Graders and adults – the influence of available orthographic information in parafoveal vision by transposing the letters of the initial trigrams of the previews (Pagán et al., 2015). Interestingly, the authors found similar effects for both groups, that is, children and adults alike were able to preprocess orthographic information. In sum, evidence suggests that 2nd Graders use parafoveal information from the second noun in a compound word pair and also benefit from phonological information presented parafoveally (Häikiö et al., 2010; Tiffin-Richards and Schroeder, 2015). Regarding the orthographic aspect of parafoveal preprocessing, however, it is still unclear whether the transposed letter manipulation induced preview costs on its own and hence resulted in an overestimation of the preview benefit (as demonstrated in Marx et al., 2015, for same-shape/different-letter masks).

## The Association of Reading Fluency, Phonological Decoding, and Rapid Naming with Parafoveal Preprocessing during Reading

In addition to the development of parafoveal preprocessing, we were interested how the capability of using parafoveal information for subsequent foveal word recognition relates to reading fluency and the children's performance in reading-related tasks. We therefore assessed the relationship between the children's reading rate in the present sentence reading task and the estimated gain of parafoveal preprocessing. Additionally, we assessed the relationship between parafoveal preprocessing and the performance of reading lists of (unrelated) words and pseudowords. Reading pseudowords taps into the children's efficiency of phonological decoding. The German orthography is very regular, that is, the grapheme–phoneme correspondence is highly consistent (in contrast to the irregular English orthography). Evidence suggests that the gain in reading fluency of children learning to read a regular orthography is primarily due to a more efficient phonological (i.e., sublexical) decoding than due to the emergence of lexical processing (i.e., whole-word recognition; Wimmer, 1993; Rau et al., 2014; Gagl et al., 2015; see Ziegler and Goswami, 2005, for a theoretical account). Thus, it will be of interest how the children's individual performance in the pseudoword reading task relates to their capability of parafoveal preprocessing.

Furthermore, we were interested in the relationship between rapid naming (RN) and the preview benefit. In RN tasks, participants are instructed to quickly and accurately name “simple” stimuli, such as objects, digits, or letters. The items are usually arranged in several lines over a page (and thus allowing for parafoveal preprocessing). A wealth of studies reported a correlation between RN and reading performance (e.g., Wolf,

1991; Wolf et al., 2000; Norton and Wolf, 2012). Expectedly, RN is considerably slower in younger readers than in older and more experienced readers. One probable cause for this speed difference in RN could be that the more experienced readers benefit from parafoveal information, whereas the younger readers do so to a much reduced extent. As yet, a direct demonstration of such a relationship is not available. Pertinent evidence, however, was provided by a recent eye movement study which demonstrated that normally developing (Chinese) readers extract information from the parafoveal items in RN, whereas in impaired (i.e., dyslexic) readers parafoveal preprocessing was markedly limited (Pan et al., 2013; see also Jones et al., 2008). A possible explanation for a relationship between parafoveal preprocessing during reading and RN (of digits) is that the increasing automaticity in processing of these (highly overlearned) symbols frees attentional resources which, in turn, can be devoted to the preprocessing of the next (i.e., parafoveal) item. Finally, an additional task assessed visual attention without the requirement of verbal processing. To be specific, we used a child-friendly adaptation of the d2-task (Brickenkamp et al., 2010) which assesses general processing speed, the efficiency of allocating visual attention and visual discrimination.

To sum up, the present eye movement study investigated parafoveal preprocessing during oral sentence reading in children of Grade 2 (with about 1 year of reading experience), Grade 4 (~3 years) and Grade 6 (~5 years). We obtained the estimates of the extent of parafoveal preprocessing by means of the novel incremental boundary paradigm (Marx et al., 2015). Our main objective was to assess the developmental course of the preview benefit. In particular, we were interested whether 2nd Grade readers already exhibit beneficial effects of parafoveal preprocessing. Additionally, we assessed how the children's reading fluency, their efficiency of phonological decoding (i.e., pseudoword reading) and their performance in RN relates to the extent of parafoveal preprocessing during reading.

## MATERIALS AND METHODS

### Participants

A total of 92 children with normal or corrected-to-normal vision participated in the study. Pupils were recruited from five different schools (from the city of Salzburg and the surrounding area). We obtained parental consent and – on the day of testing – children agreed to participate. For participation, the children received a small gift (e.g., a small ball, soap bubbles). The initial sample contained 31, 30, and 31 children from Grade 2, 4, and 6, respectively. In the present study, we were interested in normal reading development. Thus, children with a below-average and above-average reading speed – defined as a reading quotient of less than 70 ( $n = 1$ ) or more than 130 ( $n = 4$ ; see below) – were excluded from any further analysis. One additional child was excluded from the analysis due to massive data loss in the eye tracking task. The final sample consisted of 29 2nd Graders (15 females; 25 right hander; 5 children had migration background and were bilingual; age:  $8;5$  y;  $m$ ,  $SD = 0;5$ ), 27 4th Graders (13 females; 27 right hander; 7 bilinguals with migration background;

age:  $M = 10;4$ ,  $SD = 0;6$ ) and 30 6th Graders (17 females; 30 right hander; 12 bilinguals; age:  $M = 12;6$ ,  $SD = 0;6$ ). The children with a monolingual and bilingual background were comparable in their reading performance, as indexed by the reading speed test (see below; group comparison:  $t < 1$ ).

The experiment was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and it was approved by the local ethics committee of the University of Salzburg ("Ethikkommission der Universität Salzburg").

## Material

### Reading Fluency

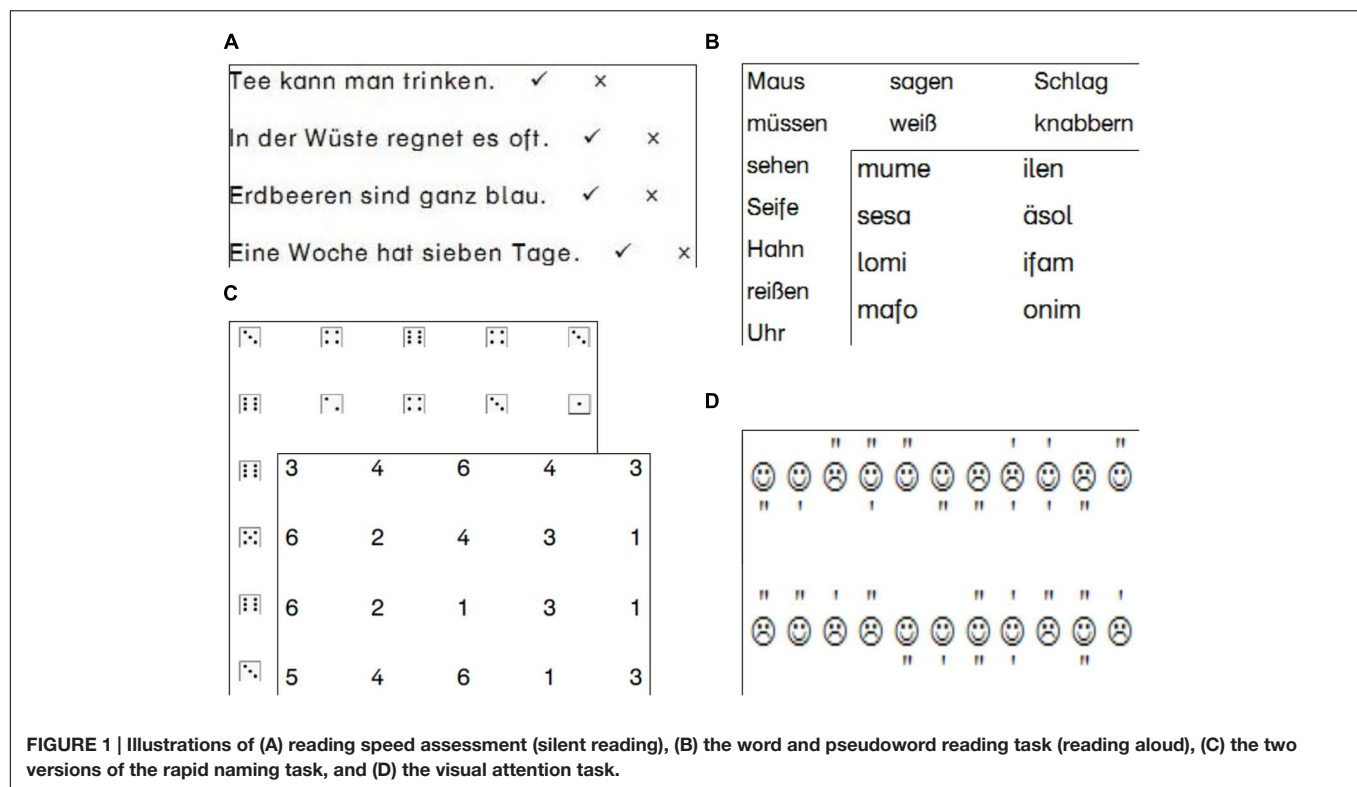
All children conducted a paper-pencil reading speed test. We used the *Salzburger Lese-Screening SLS* [Salzburg Reading-Screening] (Grade 2 and 4: SLS 1–4; Landerl et al., 1997; Grade 6: SLS 5–8; Auer et al., 2005; see **Figure 1A** for an illustration). These tests presented (age-adequate) lists of sentences which either conveys facts of basic knowledge (e.g., "A week has 7 days") or violations of basic knowledge (e.g., "Strawberries are blue"). The task of the children was to read the sentences silently and to mark each sentence as correct or incorrect within a time-limit of 3 min. As evident from the examples shown above, the decision as to the "correctness" of the sentences was easy and hence the number of correctly marked sentences is a measure of reading speed. Task performance can be expressed as a reading quotient ( $M = 100$ ,  $SD = 15$ ) based on age-norms from large-sized norming samples. In addition, we conducted a subtest of the *Salzburger Lese- und Rechtschreib-Test* [Salzburg Reading and Spelling Test] (SLRT II; Moll and Landerl, 2010; see **Figure 1B** for an illustration). The subtest required reading aloud words and pseudowords. The measure was the number of correctly read words and pseudowords within a time limit of 1 min.

### Rapid Naming

For assessing RN ability, we conducted two variants of the RN task. One presented numerals from 1 to 6; the other presented the respective dice faces (see **Figure 1C**). Each RN task consisted of 50 items in a 5-column by 10-row matrix. All RN stimuli were listed in random order with the constraint that adjacent items were not the same. Numerals were presented in an Austrian schoolbook font (20 point). Dices were presented in the same size. The children were familiarized with the test with a short practice array (two rows by five columns per stimulus type). They were timed with a stopwatch while naming the items aloud. The time was then converted to an items-per-minute measure.

### Visual Attention

The visual attention task we used (i.e., the "Smiley task") was modeled on the *d2-R* test (Brickenkamp et al., 2010). In the original version, participants are required to mark "d's" which were adorned with two quotation marks, but have to discard similar letters (e.g., "p's") with two quotation marks or "d's" with only one quotation mark. In our more child-friendly version, the letters were replaced by line-drawings of happy and unhappy faces (i.e., "smileys" and "frownies"; see **Figure 1D**). Children had to mark the smiley faces adorned with two quotation



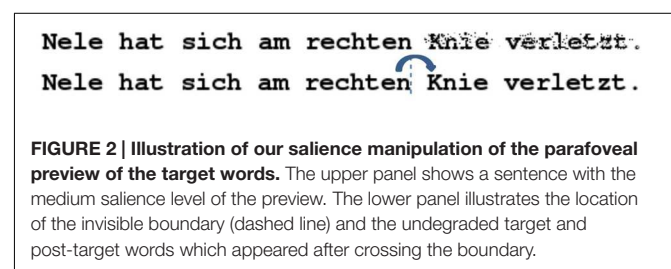
marks. Distractors were smileys and frownies with less or more quotation marks (frownies with 2 marks also served as distractor items). Items were presented in nine lines with 47 smiley faces in each line (30 smiley faces; with an average of 20 right choices) per line. For each line, the children had 20 s whereupon they had to stop and start with the next line. We considered the mean number of correctly marked smiley faces within 1 min as our measure of attention (more specifically, the test assesses general processing speed, serial allocation of visual attention and visual discrimination).

### Eye Tracking Task

For the eye tracking task we presented 90 sentences in which we embedded one target word per sentence (i.e.,  $N = 30$  sentences for Grade 2 children;  $N = 60$  sentences for Grade 4 children;  $N = 90$  sentences for Grade 6 children). The target words were exclusively nouns and had a mean length of five letters (range: 4–6 letters) and a mean frequency (occurrences per million) of 105 according to the SUBTLEX-DE norms (Brysbaert et al., 2011). Note that we used the same sentences as in a previous study from our lab (Marx et al., 2015). The target words were – according to a Latin square design – rotated between the three salience conditions for each Grade. Sentences were constructed in such a way that at least three words preceded and at least one word followed the target word ( $M = 5.4$  and  $2.5$ , respectively). The pretarget words were of medium-length and (on average) high-frequency adjectives. Specifically, the mean length of the pretarget word was 5.26 letters ( $SD = 0.84$ ; range: 4–8) and the mean frequency of their lemma-form (i.e., the uninflected

form of the word) was 204 per million (word-form:  $M = 85$  per million). The length of the experimental sentences ranged from 6 to 12 words ( $M = 8.84$ ,  $SD = 1.11$ ). The sentences were typed in a bold and mono-spaced font. Each character had a width of 8 pixels on the display screen (whose specifications are provided in the “Apparatus” section). From a viewing distance of 50 cm a single character had a width of  $\sim 0.4^\circ$  of visual angle.

The salience manipulation (i.e., visual degradation) of the stimuli was administered by using the *pixmap*-package (Bivand et al., 2008) and an in-house *R-script*. We had three preview conditions (i.e., parafoveal salience manipulations). In each preview condition all letters of the target and all words thereafter were degraded, that is, a certain amount of black pixels was displaced. The amount of displaced pixels were 0, 10, and 20% for our three levels of degradation (henceforth, we refer to the levels as high, medium, and low salience). An example sentence of our experimental set-up is shown in Figure 2.





## Procedure

### Psychometric Assessment

At first, we administered the reading speed test in the children's classrooms. The further psychometric assessments (as well as the eye tracking experiment) were conducted over 2 days during which children were seen individually in a quiet area detached from the classroom. On a rotating basis, first and second day procedure and order of tasks were counterbalanced across participants, whereby the psychometric measures lasted approximately 40 min and the eye tracking task lasted approximately 20 min.

### Eye Tracking

First, we performed a horizontal 3-point calibration routine to familiarize the children with calibrating the eye tracking system. This routine was repeated until the child achieved an average tracking error below  $0.5^\circ$  of visual angle. Then, five familiarization trials for the sentence reading task were administered after which the calibration was repeated – now with a more stringent criterion (average tracking error  $< 0.3^\circ$ ). Then, we presented the 30, 60, or 90 experimental sentences (dependent on Grade; see section “Eye Tracking Task”). A trial started with a *fixation check*, that is, the presentation of a fixation cross at the left side of the screen (vertically centered). Calibration was repeated when the fixation check failed (but not later than the presentation of 20, 35, or 50 sentences for Grade 2, 4, and 6, respectively). When the system detected a fixation on the fixation cross, the sentence was presented. Display changes were realized with the invisible boundary technique (Rayner, 1975). The boundary was placed at the very end of the pretarget word. Crossing the boundary triggered the presentation of the identical target (and post-target) word(s) – in cases where a high salience preview was presented – or the unmutated target (and post-target) word(s) – in cases where a medium or low salience preview was presented. The children read the sentences aloud. The experimenter noted reading errors (mostly minor misarticulations, such as, e.g., improper lengthening or shortening of vowels with frequent immediate self-correction by the children).

### Apparatus

Eye movements were recorded monocular for the right eye with a sampling rate of 500 Hz with an EyeLink 1000 (SR Research, Canada). We used the Desktop mount configuration with the “remote” setup which compensates for head movements (by tracking a target sticker on the children's forehead). The children sat at a viewing distance of approximately 50 cm to the 17 inch CRT-monitor (640 × 480 pixel resolution with a 200 Hz frame rate).

### Eye Movement Measures

We reasoned that the effect of parafoveal preprocessing will be most evident in the initial fixation on the target words. Thus, we considered first fixation (FF) duration as our primary dependent variable. Additionally, we report single fixation duration (SF; i.e., when target words were processed with a SF) and gaze duration (i.e., the sum of all fixations on a target word during

first pass reading). Furthermore, we report the initial landing position (ILP), that is, the location of the FF on the target words.

## Data Treatment and Analysis

In total, we administered 5,190 trials (i.e., 29, 27, and 30 children from Grade 2, 4, and 6 read 30, 60, and 90 sentences, respectively; see above). After removal of trials with data loss and outlying fixation times on the target words, 3,860 and 3,851 trials remained for the analysis of FF and gaze duration, respectively. The criteria for outliers were fixations times shorter than 80 ms and longer than 2.65 standard deviation above the individual mean of the participant. For the analysis of SF, we only obtained a total of 1,938 trials, because children seldom processed a word with a SF (see “Results” section). Eye movement data were analyzed by means of linear mixed effects (LMM) modeling using the *lmer*-function of the *lme4*-package (Bates et al., 2015) running within the R environment for statistical computing (R Core Team, 2015). For our global eye movement measures we considered each word except the sentence-initial word and the target word (whose parafoveal preview was manipulated). The model assessed – as fixed effect – the linear effect of Grade and accounted for the random effects of subjects (i.e., the individual children) and items (i.e., the target words). The syntax for this model was *measure ~ grade + (1 | subject) + (1 | item)*. For the analyses of the experimental effect of our salience manipulation on FF, SF, and gaze duration we used a more sophisticated model specification whose syntax was as follows: *measure ~ salience + grade + salience:grade + (1 + salience + grade + salience:grade | subject) + (1 | item)*. The model examined – as fixed effect – the linear effects of Grade and salience and the two-way-interactions between these effects. Besides these fixed effects, the model accounted for the random effects of subjects on the intercept of the model and on the slopes of the salience and Grade effects as well as for random effects of the items. Following standard convention, fixed effects were considered as significant when the corresponding *t*-value was greater than 1.96 (which corresponds to an alpha-level of  $p < 0.05$ ). We *log*-transformed FF, SF, and gaze duration (by the natural logarithm) before entering the analyses, because their distributions were right skewed (the figures, however, presents untransformed data).

## RESULTS

### Reading Rate and Psychometric Measures

Mean task performances as a function of Grade are presented in **Table 1**. The first line of the **Table 1** presents the mean reading quotient of the children from Grade 2, 4, and 6: The groups of children exhibited, on average, normal reading rates (compared to the respective age-norms of  $M = 100$  and  $SD = 15$ ). Accordingly, a univariate ANOVA revealed no group differences;  $F < 1.1$ . In absolute terms, reading rate almost doubled from Grade 2 to Grade 6 as evident from the word-per-minute measure of reading aloud lists of unrelated words;  $F(2,85) = 42$ ,  $p < 0.001$ . The gain in reading

**TABLE 1 | Means and standard deviations of the psychometric measures and global eye movements.**

	Grade 2		Grade 4		Grade 6	
	Mean	SD	Mean	SD	Mean	SD
<b>Psychometric measures</b>						
Silent reading [RQ <sup>a</sup> ]	103	11	101	12	98	16
Reading aloud words	50	12	71	22	95	21
Reading aloud p-words <sup>b</sup>	33	7	41	13	56	17
Rapid naming						
Dices	79	19	102	16	115	22
Digits	98	25	134	25	153	32
Visual attention	42	6	56	10	69	12
<b>Eye movement task</b>						
Words per minute	74	18	102	26	131	28
Reading errors [%]	23	23	10	8	6	5
N of fixations per word	1.9	0.3	1.7	0.4	1.5	0.2
Fixation duration [ms]	451	110	360	82	288	45
Saccade length [letters]	3.7	0.8	4.2	0.7	4.6	0.6
Regressions [%]	18	9	19	9	15	6

<sup>a</sup>Reading Quotient, <sup>b</sup>pseudowords.

speed was significant between each Grade (*post hoc* pairwise comparisons:  $t_s > 4.15$ ,  $p_s < 0.001$ ). Likewise, reading aloud lists of unrelated pseudowords showed an improvement with Grade;  $F(2,85) = 25$ ,  $p < 0.001$  (*post hoc* pairwise comparisons:  $t_s > 2.96$ ,  $p_s < 0.01$ ). Furthermore, the number of words-read-per-minute (assessed in our eye tracking experiment; lower section of **Table 1**) increased with Grade;  $F(2,85) = 40$ ,  $p < 0.001$ . Pairwise comparisons revealed that the difference was significant between each Grade ( $t_s > 3.98$ ,  $p_s < 0.001$ ). Likewise, the reading accuracy (assessed in our eye tracking experiment) improved with Grade; Kruskal–Wallis  $\chi^2 = 17.48$ ,  $p < 0.001$ . The differences were significant between each Grade (Mann–Whitney  $U_s < 265$ ,  $p < 0.04$ ). Furthermore, children became faster in both versions of the RN task; main effect of Grade:  $F(2,83) = 31.90$ ,  $p < 0.001$ . Improvements – for both versions of the task – were evident between all Grades;  $t_s > 2.55$ ,  $p_s < 0.02$ . With regard to differences between the RN versions, the children’s performance was faster for the digit version than for the dice version; main effect of RN version:  $F(1,83) = 241$ ,  $p < 0.001$ . This difference was more pronounced in Grade 4 and 6 than in Grade 2; Grade by RN version:  $F(2,83) = 9.54$ ,  $p < 0.001$ . For our measure of visual attention (i.e., the “Smiley task”), we observed a continuous improvement with Grade;  $F(2,85) = 55$ ,  $p < 0.001$  (pairwise comparisons: all  $t_s > 4.3$ ,  $p_s < 0.001$ ).

## Global Eye Movement Measures

As evident from the lower section of **Table 1**, the mean number of fixations per word decreased with Grade;  $b = -0.207$ ,  $SE = 0.040$ ,  $t = -5.13$ . This reduction was significant between both, Grade 2 and 4;  $b = -0.240$ ,  $SE = 0.093$ ,  $t = -2.58$ , and Grade 4 and 6;  $b = -0.175$ ,  $SE = 0.079$ ,  $t = -2.20$ . The mean fixation duration decreased with Grade;  $b = -0.184$ ,  $SE = 0.024$ ,  $t = -7.76$ , and

the difference was significant between Grade 2 and 4, and Grade 4 and 6;  $b = -0.189$ ,  $SE = 0.053$ ,  $t = -3.54$  and  $b = -0.180$ ,  $SE = 0.048$ ,  $t = -3.77$ , respectively. The mean forward saccade length increased with Grade;  $b = 0.462$ ,  $SE = 0.087$ ,  $t = 5.30$  ( $b = 0.468$ ,  $SE = 0.192$ ,  $t = 2.43$  and  $b = 0.458$ ,  $SE = 0.163$ ,  $t = 2.81$  for the Grade 2 – 4 and 4 – 6 comparisons). Finally, there was a linear, but insignificant trend toward fewer regressions with Grade;  $b = -0.019$ ,  $SE = 0.010$ ,  $t = -1.82$ .

## Target Words

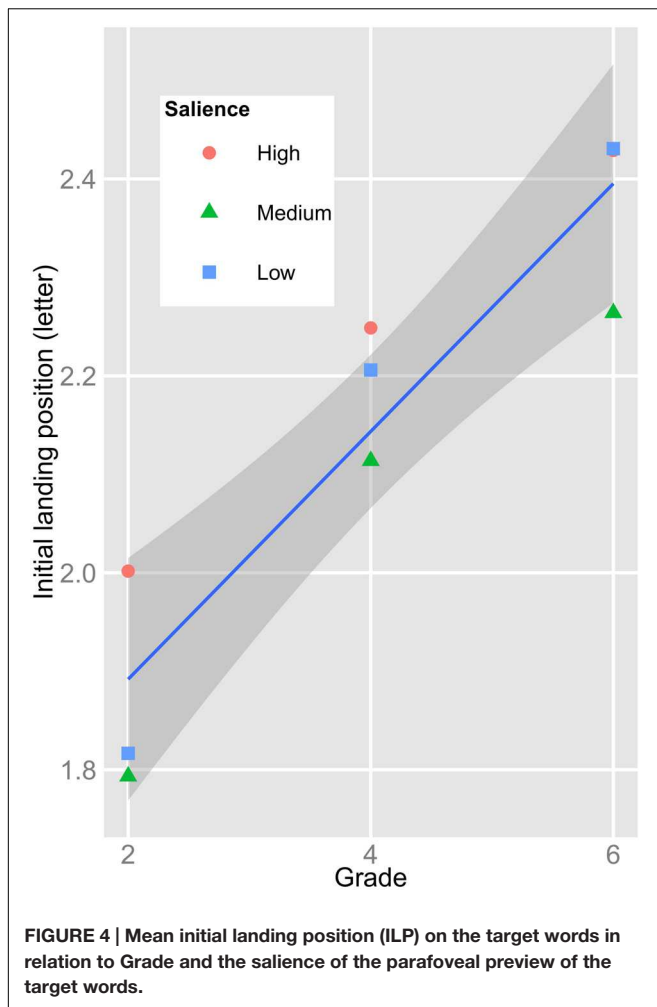
The target words were rarely skipped ( $M < 3.6\%$  for each Grade) and seldom processed with a SF, i.e., in only 12, 20, and 28% of the trials for Grade 2, 4, and 6, respectively. **Figure 3** presents fixation time measures on the target words in relation to our salience manipulation of the target words’ parafoveal preview and Grade. As evident from **Figure 3**, fixation durations became progressively shorter with Grade. This development toward shorter fixation durations was reflected by a main effect of Grade (see **Table 2** for model estimates and the corresponding  $t$ -values). Critically, all Grades exhibited shorter FF durations for high-salience than for low-salience previews of the target words. For the undegraded (high-salience) previews, the means of FF were 526 ms ( $SD = 112$  ms), 349 ms ( $SD = 79$  ms), and 294 ms ( $SD = 37$  ms) for Grade 2, 4, and 6, respectively. For the low-salience previews, the means were 571 ( $SD = 134$ ), 408 ( $SD = 54$ ), and 338 ms ( $SD = 39$ ) resulting in mean differences of 45, 59, and 44 ms for Grades 2, 4, and 6, respectively. Accordingly, the LMM revealed a main effect of salience but the interaction between salience and Grade was not significant (see **Table 2**).

Remember that the children seldom processed the words with a SF and, thus, the analysis of SF duration should not be overrated. In short, **Figure 3** shows that Grade 4 and Grade 6 exhibited shorter SF durations for high-salience than for low-salience previews of the target words. The children from Grade 2 did not exhibit such an effect. Accordingly, the LMM revealed an interaction between salience and Grade; the main effect of salience did not reach significance. Separate LMMs for each Grade revealed significant effect of salience in each Grade. The fixed effects of salience, however, were much higher for the children of Grade 4 ( $b = 0.111$ ,  $SE = 0.027$ ,  $t = 4.18$ ) and Grade 6 ( $b = 0.134$ ,  $SE = 0.108$ ,  $t = 12.42$ ) than for the children of Grade 2 ( $b = 0.073$ ,  $SE = 0.033$ ,  $t = 2.20$ ). It is noteworthy that – as evident from **Figure 3** – SF were, on average, longer than FF (see “Discussion”). Pairwise comparisons (independent of the level of salience) revealed that this difference was significant for each Grade; all  $t_s > 5.1$  ( $df = 19$ , 26, and 29 for Grade 2, 4, and 6, respectively), all  $p_s < 0.001$ .

The LMM for gaze duration did not reveal a significant main effect of salience, but a significant interaction between salience and Grade. Separate models revealed that the fixed effect of salience was significant in Grade 4 ( $b = 0.065$ ,  $SE = 0.021$ ,  $t = 3.10$ ) and Grade 6 ( $b = 0.098$ ,  $SE = 0.015$ ,  $t = 6.55$ ). For Grade 2, the effect of salience did not reach significance ( $b = 0.052$ ,  $SE = 0.028$ ,  $t = 1.85$ ).

**Figure 4** presents the ILP of the children in relation to Grade and the salience of the parafoveal preview. As evident from

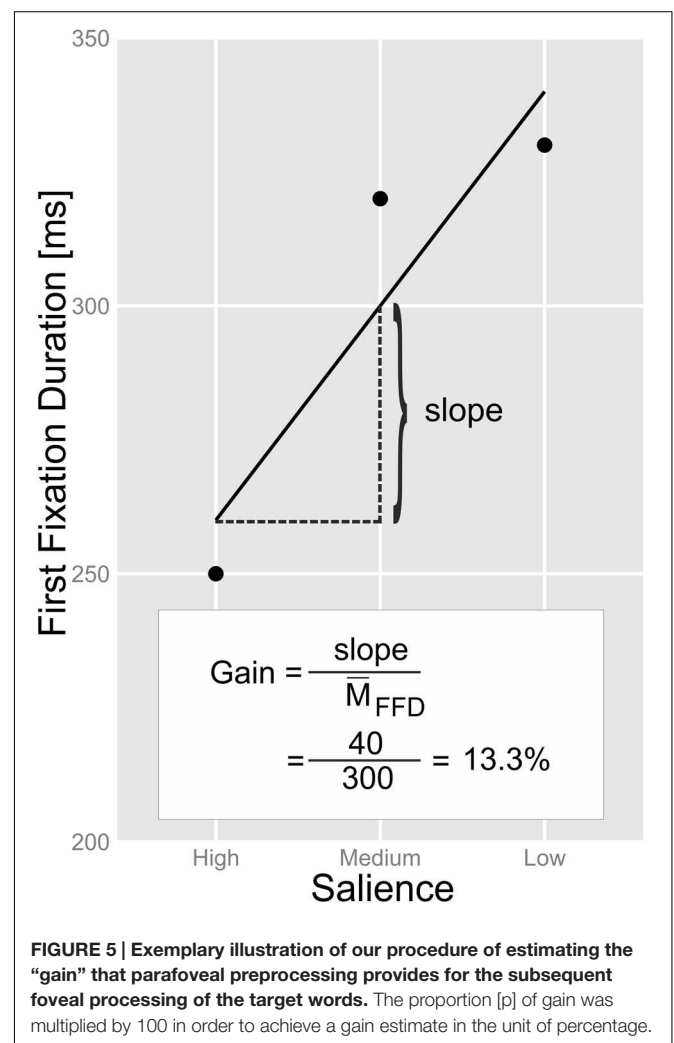




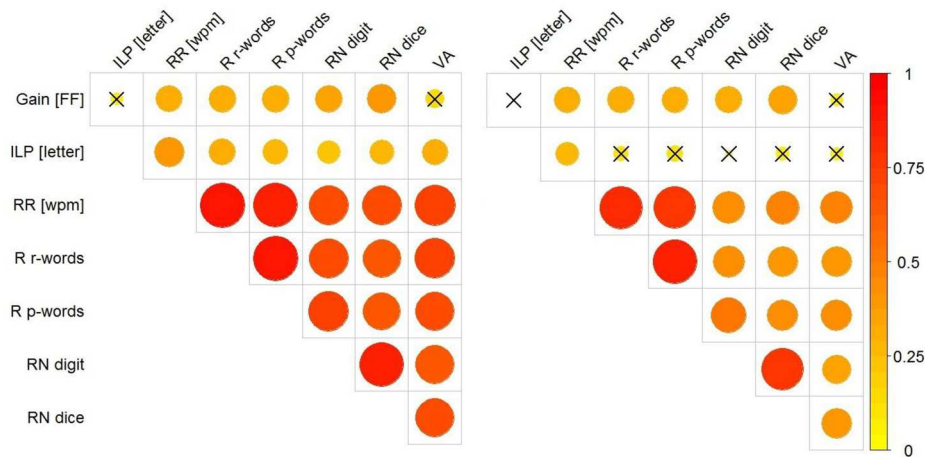
estimate of the magnitude of the preview benefit; see **Figure 5**), (ii) the ILP on the target words, (iii) the reading rate as expressed by words-per-minute (from the eye tracking experiment), the rate of reading aloud columns of (iv) words and (v) pseudowords, the performance in (vi) the dice-version and (vii) the digit-version of the RN task, and (viii) the performance in the visual attention task (VA; i.e., the “Smiley” variant of the d2-test). The left panel of **Figure 6** shows the correlations for all participants irrespective of Grade; the right panel shows the correlations when we partialled-out the effect of Grade. Contrasting full versus partial correlations gives us indications as to whether an association of a variable with our index of parafoveal preprocessing (i.e., “gain”) reflects “merely” a Grade-related improvement in both measures or whether there is a specific (Grade-independent) relationship. As evident from the left panel of **Figure 7**, the reading rate measures and the performance in the two versions of the RN task were highly (inter-)correlated. Furthermore, reading rates and RN were highly correlated with the performance in the VA task. Partialling-out Grade reduced the size of the correlations of RN and VA with the reading rate measures. Critically, our estimate of the usage of parafoveal information for subsequent foveal word recognition (gain) correlated (moderately) with our

various reading rate measures and with RN (see also **Figure 7**). These correlations were significant even when we partialled-out the effect of Grade. The gain due to parafoveal preprocessing was not correlated with the ILP on the target words and not with VA. ILP was reliably associated with reading rate during reading the experimental sentences (i.e., from the eye movement assessment). This association remained significant when we partialled-out Grade. The correlations between ILP and the reading rates for words and pseudowords, RN and VA were insignificant after controlling for Grade.

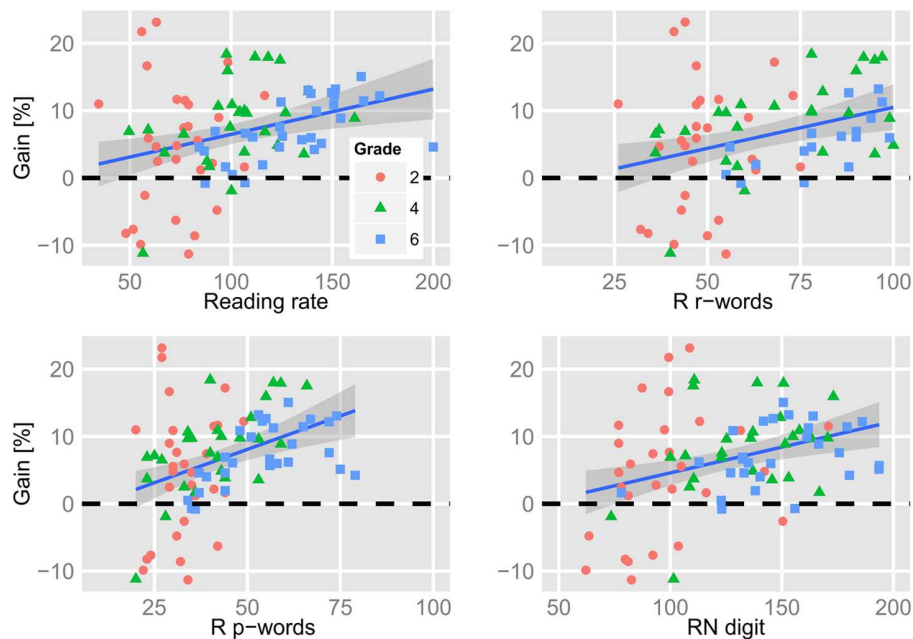
**Figure 7** shows the relationship between the estimated gain due to parafoveal preprocessing and selected psychometric measures (with the individual scores of the participating children). From the top-right corner to the bottom-left corner, **Figure 7** shows how the gain measure relates to the reading rate from the eye tracking/sentence reading task, the word-list reading task, the pseudoword-list reading task and the RN-digit task. Average reading rates were task-dependent (RN digits > reading words in sentences > reading list of words > reading pseudowords; see **Table 1**). **Figure 7** shows







**FIGURE 6 |** Correlations (Pearson's  $r$ ) between the proportional reduction of first fixation duration in relation to the salience of the parafoveal preview (Gain [FF]), the ILP on the target words (ILP), the word-per-minute (wpm) rate of reading aloud sentences (RR; obtained from the eye tracking task), the wpm-rate of reading aloud lists of words (R r-words) and pseudowords (R p-words), the items-per-minute measure of the two versions of the rapid naming task (i.e., RN digits and dice faces) and the performance in the visual attention task (VA). The (Left) panel presents the correlations irrespective of the Grade-level of the children; the (Right) panel shows the correlations after partialling-out Grade. The size of the correlations is represented by the size (and the color) of the circles; correlations of  $r > 0.23$  were significant ( $p < 0.05$ ); insignificant correlations are marked with an X. For creating this Figure, we used the *corrplot*-package (Wei, 2013).



**FIGURE 7 |** Relation between the percentage of gain (preview benefit) with psychometric measures (reading rate from the eye tracking task, reading real words, reading pseudowords, and rapid naming digits). Positive values (i.e., values above the bold dashed line) indicate gain that is the proportional reduction of the first fixation duration in relation to the salience of the parafoveal preview of the target words (see Figure 5 and main text for details).

that we observed stable gains due to parafoveal preprocessing when the children read more than 100 words-per-minute in the sentence reading task. For reading lists of words and pseudowords, the respective figures were  $\sim 75$  and  $\sim 50$  items-per-minute. For RN of digits, we observed relatively stable gains when the children's rate was greater than 150 items-per-minute.

Finally, we assessed which of the four rate measures is the most potent predictor of the ability of gaining parafoveal information from the upcoming word. To this end, we fitted a linear model with the four predictors and submitted this model to the *stepAIC*-function of the *MASS*-package (Venables and Ripley, 2002). This function performs a stepwise model selection on the basis of the

Akaike information criterion (AIC). This analysis revealed that the performance in the pseudoword list reading task is the best predictor of the preview benefit in our sample of German-reading children.

## DISCUSSION

The main objective of the present developmental eye tracking study was to examine when children begin to effectively utilize parafoveal information during reading. In an earlier study from our lab (Marx et al., 2015), we found that children with about 3 years of reading experience (i.e., children in Grade 4 of primary school) exhibited a substantial preview benefit – similar to children with about 5 years of reading experience (Grade 6). Thus, we assumed that parafoveal preprocessing emerges early during reading acquisition. In the present study, we tested children from Grades 2, 4, and 6.

For the assessment of the magnitude of parafoveal preprocessing we used a recently developed paradigm which combines the classical invisible boundary paradigm (Rayner, 1975) from the field of eye movement research with the rationale of the *incremental priming* technique (Jacobs et al., 1995) from the field of visual word recognition. The rationale behind administering this novel technique was that recent evidence indicated that the application of parafoveal masks – which is the traditional approach for estimating the preview benefit in the context of the invisible boundary paradigm – may lead to an overestimation of the preview benefit (Hutzler et al., 2013; Kliegl et al., 2013; Marx et al., 2015). The *incremental boundary* technique (which systematically manipulates the salience of the parafoveal preview of the target words; see “Introduction”) is much less susceptible to such a bias (Marx et al., 2015).

The main finding of the present study is that children from Grades 2, 4 and 6 exhibited substantially shorter FF durations with increasing salience of the parafoveal preview, that is, they exhibited a preview benefit. For FF duration on the target words, the incremental boundary approach (i.e., comparing mean FF duration for high-salience with those of low-salience previews) revealed estimates of the size of the preview benefit of about 45 ms for children of Grades 2 and 6 and for children of Grade 4 the size was about 60 ms. These figures translate to a shortening of fixation duration of about 8% in Grade 2 and of about 15 and 13% in Grade 4 and 6 when preprocessing of a valid (i.e., high salience) preview is possible compared to instances in which parafoveal preprocessing is hindered (by a visually degraded preview). Thus, we found clear evidence of a parafoveal preview benefit on FF duration for all of the Grades.

The instances in which the children processed the words with a SF were rare – even in the most experienced readers of Grade 6 (<30%). The low number of SF cases indicates that our children (learning to read the regular German orthography) achieve visual word recognition primarily due to serial (grapheme–phoneme) decoding – even when they already have considerable reading experience (for similar results and

interpretation see Rau et al., 2014 and Gagl et al., 2015). For the children of Grades 4 and 6, however, we observed a preview benefit on SF duration (i.e., reliable effects of our manipulation of the parafoveal preview). Replicating previous findings (e.g., Hawelka et al., 2010), the mean duration of SF were longer than the average duration of FF on the target words. Processing words with a SF has been considered as reflecting whole-word recognition and the prolongation of SF in comparison to FF may reflect the completion of lexical processing, that is, accessing whole-word phonology and word meaning (Hawelka et al., 2010). Thus, parafoveal preprocessing seems to be beneficial for whole-word recognition even if this manner of word recognition is still comparatively rare (as indicated by the small proportion of singly fixated words).

In addition to fixation times, we assessed the ILP on the target words in relation to the Grade-level of the children and to the salience of the preview. The motivation for including this measure was twofold. First, we were interested in the development of the visual scanning behavior during the initial years of reading acquisition. Second, we were interested in the relationship between the extent of parafoveal preprocessing (indexed by the size of the preview benefit) and the saccadic targeting of the upcoming word. With regard to the first aspect, experiments using single word presentation (with French children) revealed that beginning readers quickly acquire an adult-like tendency to fixate at the *optimal viewing position*, that is, (slightly left of) the word center (Aghababian and Nazir, 2000). This shift in targeting the center of a word – as opposed to targeting a word's initial letters – was previously attributed to the progress from laborious sublexical grapheme–phoneme conversion toward more efficient whole-word recognition (MacKeben et al., 2004; Hawelka et al., 2010; Rau et al., 2015). The efficiency of processing a word by means of sublexical decoding, however, is supposed to be largely dependent on the orthographic depth of the to-be-learned language. To illustrate, a recent eye movement study, which directly compared sentence reading in German (a shallow orthography) and English (a deep orthography; Seymour et al., 2003; Share, 2008) – showed that the German readers relied more on small-unit decoding than their English peers (Rau et al., 2015). Supporting the notion of such a small-unit decoding strategy, recent eye movement studies in regular orthographies reported that beginning readers tend to aim the incoming saccade at the word beginning (Gagl et al., 2015). To illustrate, Gagl et al. (2015) reported – for an experiment with single word presentation – that German-reading children of Grade 2 and 4 fixated the word beginning with little influence of word length on initial fixation location. Likewise, in our previous study (Marx et al., 2015), we found that the ILP of children of Grades 4 and 6 (in a sentence reading task with valid and invalid previews of target words) was at the beginning of the target words. In the present study, we found a significant developmental trend of initial fixation location toward the word center. Furthermore, the ILP was reliably correlated with the reading rate (even when the Grade-level was partialled-out). The size of the Grade effect, however, was – in absolute terms – small (half a letter from Grade 2 to Grade 6). Thus, our

finding conforms to the notion that the progress from grapheme–phoneme conversion toward whole-word recognition proceeds slowly in regular orthographies.

With regard to our second interest, we found no association of the ILP with the extent of parafoveal preprocessing. This was even the case, when we correlated these two measures irrespective of Grade (i.e., without partialling-out Grade). The absence of an association between the preview benefit and the ILP conforms to the assumed decoupling of oculomotor control and visual attention (as it is implemented, for example, in the E-Z Reader model of eye movement control during reading; Reichle et al., 1998, 2003). In the conceptualization of the E-Z reader model, the processing of the length of the next word is considered as a basal visual, pre-attentive process. Accordingly, we did not find an association of the ILP with our measure of visual attention (i.e., our variant of the d2-test which assesses the serial allocation of visual attention and visual discrimination) after accounting for Grade-level effects (i.e., after partialling-out age-related improvement in visual attention). The mechanism that oculomotor control and visual attention operates independently can explain the fact that mature readers frequently skip words during reading. The fact that we did not find an association between the amount of parafoveal preprocessing and saccadic aiming in beginning readers could indicate that the functional separation of visual attention and oculomotor planning is already in place during reading acquisition (when word skippings are still very rare).

## The Association of the Parafoveal Preview Benefit with Rapid Naming and Pseudoword Reading

We administered two versions of the RN task, that is, a “standard” version which required the naming of digits (ranging from 1 to 6) and an equivalent version in which we substituted the Arabic numerals with the corresponding dice faces. The rationale for the administration of these two versions was that we assumed that the children of Grade 2 (with only about 1 year of formal education) might not yet exhibit automaticity in processing (in future over-learned) orthographic representations (i.e., the Arabic numerals). The ensuing expectations were that (i) the children from Grade 2 would exhibit a more similar performance in the two RN versions, whereas the older children would perform better in the digit version and (ii) that the association of RN of digits may become stronger with increasing Grade-level. A recent eye movements study by Pan et al. (2013) indeed showed that the eye-voice span is larger during RN of digits than during RN of dice faces (which was interpreted as reflecting the higher automaticity of processing Arabic numerals). Moreover, this effect was markedly more pronounced in typically developing readers than in dyslexic readers (indicating a less automatized processing of Arabic numerals in the latter group). The present findings conform to the notion of heightened automaticity for over-learned orthographic symbols. In each Grade, the children performed better in the digit-version than in the dice-version of the RN task, but the difference was more pronounced in the higher Grades. However, the performance in both versions was

associated equally with our estimate of parafoveal preprocessing and this association did not depend on reading experience (i.e., Grade level). The similar association of RN dice faces and digits with parafoveal preprocessing may reflect the shared requirement of coordinating the serial allocation of visual attention (in the direction of reading) and accessing a phonological representation as figured by the visual scanning hypothesis of the relationship of RN with reading (e.g., Kuperman et al., 2016).

The best predictor of parafoveal preprocessing was the children’s performance in the pseudoword reading task. The task assessed the children’s efficiency of phonological decoding. As aforementioned, the developmental transition from sublexical decoding to (lexical) whole-word recognition seems to be a slow process in regular orthographies (e.g., Rau et al., 2014) and hence the improvement in reading rate with increasing experience is – at least partly – due to a gain in the efficiency of phonological decoding (e.g., Wimmer, 1993; Gagl et al., 2015). The present finding adds to this notion by showing that children who excelled on the pseudoword reading task exhibited the largest preview benefit.

## Limitations and Future Directions

One could conceive the present study’s requirement of reading aloud as a limitation for studying the development of the preview benefit, because reading aloud may reduce the extent to which readers engage in parafoveal preprocessing. To illustrate, Ashby et al. (2012) found – in adult participants – that the preview benefit is diminished in oral reading compared to silent reading. However, silent reading is unusual for children – particularly in the early years of primary schools. Another limiting issue, one could argue, is the high variance in the performance of the children from Grade 2. The variance in our dependent measures was much lower in Grade 4 and 6. This pattern conforms to the prediction of, for example, the rate-amount model (Faust et al., 1999) that increasing average efficiency is accompanied with a reduction in variance. To account for a global factor such as general processing speed (e.g., Zoccolotti et al., 2008) was, however, beyond the scope of the present study.

With regard to future directions, the present study (together with the Marx et al., 2015 study) showed that the incremental boundary technique is an adequate tool for studying the emergence and the development of parafoveal preprocessing in developing readers. Future studies may apply the technique to study further aspects of parafoveal preprocessing (for which the evidence is, as yet, based primarily on samples of adult readers). Such aspects are, for example, the relative importance of a word’s initial versus its final letters for parafoveal preprocessing (e.g., Briehl and Inhoff, 1995; Gagl et al., 2014) or the effect of foveal load on the preview benefit (Henderson and Ferreira, 1990).

## CONCLUSION

The present study provides information as to when parafoveal information is effectively utilized during oral sentence reading. Overall, the findings reveal that children with about 1 year of reading experience start to utilize parafoveal information for

subsequent foveal word recognition. However, we observed an association of the preview benefit with reading fluency (indexed by the word-per-minute reading rate) – which substantially overlapped between Grades. Thus, the individual reading competence seems to be the more important constituent of the effective use of parafoveal information for subsequent foveal word recognition than reading experience as indexed by Grade-level. The best predictor of parafoveal preprocessing in our sample of children learning to read a regular orthography was their performance in a pseudoword reading task assessing the efficiency of phonological decoding: The best decoders exhibited the greatest preview benefit.

## AUTHOR CONTRIBUTIONS

CM performed the experiment, CM and SH analyzed the data. CM, SH, and SS wrote the manuscript. CM and SS prepared the

figures, CM, FH, and SH conceived the experiment. All authors reviewed the manuscript.

## FUNDING

This work was supported by the Austrian Science Fund (FWF) under grant P25799B23.

## ACKNOWLEDGMENTS

We would like to thank the children who participated in the study and the principals and teachers who gave us permission to conduct the study in their schools. We are grateful to Ramona Zintl for her help in data collection. We thank Franziska A. Fowles for proof reading.

## REFERENCES

- Aghababian, V., and Nazir, T. A. (2000). Developing normal reading skills: aspects of the visual processes underlying word recognition. *J. Exp. Child Psychol.* 76, 123–150. doi: 10.1006/jecp.1999.2540
- Ashby, J., Yang, J., Evans, K. H. C., and Rayner, K. (2012). Eye movements and the perceptual span in silent and oral reading. *Atten. Percept. Psychophys.* 72, 634–640. doi: 10.3758/s13414-012-0277-0
- Auer, M., Gruber, G., Mayringer, H., and Wimmer, H. (2005). *Salzburger Lese-Screening für die Klassenstufen 5-8 (SLS 5-8)*. Göttingen: Verlag Hans Huber.
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). *lme4: Linear Mixed-Effects Models Using Eigen and S4.R Package Version 1.1-8*. Available at: <http://CRAN.R-project.org/package=lme4>
- Bivand, R., Leisch, F., and Mächler, M. (2008). *Pixmap: Bitmap images ("Pixel maps"; R Package, Version 0.4-9)*. Available at: <https://cran.r-project.org/package=pixmap>
- Brickenkamp, R., Schmidt-Atertz, L., and Liepmann, D. (2010). *Test d2- Revision: Aufmerksamkeits- und Konzentrationstest*. Göttingen: Hogrefe.
- Brihl, D., and Inhoff, A. W. (1995). Integrating information across fixations during reading: the use of orthographic bodies and of exterior letters. *J. Exp. Psycho. Learn. Mem. Cogn.* 21, 55–67. doi: 10.1037/0278-7393.21.1.55
- Brysbaert, M., Buchmeier, M., Conrad, M., Bölte, J., and Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Exp. Psychol.* 58, 412–424. doi: 10.1027/1618-3169/a000123
- Faust, M. E., Balota, D. A., Spieler, D. H., and Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: implications for group differences in response latency. *Psychol. Bull.* 125, 777–799. doi: 10.1037/0033-2909.125.6.777
- Gagl, B., Hawelka, S., Richlan, F., Schuster, S., and Hutzler, F. (2014). Parafoveal preprocessing in reading revisited: evidence from a novel preview manipulation. *J. Exp. Psycho. Learn. Mem. Cogn.* 40, 588–595. doi: 10.1037/a0034408
- Gagl, B., Hawelka, S., and Wimmer, H. (2015). On sources of the word length effect in young readers. *Sci. Stud. Read. Sci. Stud.* 19, 289–306. doi: 10.1080/10888438.2015.1026969
- Häikiö, T., Bertram, R., and Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: evidence from the boundary paradigm. *Q. J. Exp. Psychol.* 63, 1982–1998. doi: 10.1080/17470211003592613
- Häikiö, T., Bertram, R., Hyönä, J., and Nieme, P. (2009). Development of letter identity span in reading: evidence from the eye movement moving window paradigm. *J. Exp. Child Psychol.* 102, 167–181. doi: 10.1016/j.jecp.2008.04.002
- Hawelka, S., Gagl, B., and Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition* 115, 367–379. doi: 10.1016/j.cognition.2009.11.004
- Henderson, J. M., and Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: implications for attention and eye movement control. *J. Exp. Psycho. Learn. Mem. Cogn.* 16, 417–429. doi: 10.1037/0278-7393.16.3.417
- Hutzler, F., Fuchs, I., Gagl, B., Schuster, S., Richlan, F., Braun, M., et al. (2013). Parafoveal X-masks interfere with foveal word recognition: evidence from fixation related brain potentials. *Front. Syst. Neurosci.* 7:33. doi: 10.3389/fnsys.2013.00033
- Jacobs, A. M., Grainger, J., and Ferrand, L. (1995). The incremental priming technique: a method for determining within-condition priming effects. *Percept. Psychophys.* 57, 1101–1110. doi: 10.3758/BF03208367
- Jones, M. W., Obregon, M., Kelly, M. L., and Branigan, H. P. (2008). Elucidating the component processes involved in dyslexic and non-dyslexic reading fluency: an eye-tracking study. *Cognition* 109, 389–407. doi: 10.1016/j.cognition.2008.10.005
- Kliegl, R., Hohenstein, S., Yan, M., and McDonald, S. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Q. J. Exp. Psychol.* 66, 581–600. doi: 10.1080/17470218.2012.658073
- Kuperman, V., Van Dyke, J. A., and Regina Henry, R. (2016). Eye-movement control in RAN and reading. *Sci. Stud. Read* 20, 173–188. doi: 10.1080/10888438.2015.1128435
- Landerl, K., Wimmer, H., and Moser, E. (1997). *Salzburger Lese- und Rechtschreibtest: Verfahren zur Differentialdiagnose von Störungen des Lesens und Schreibens für die 1. und 4. Schulstufe*. Bern: Huber.
- MacKeben, M., Trauzettel-Klosinski, S., Reinhard, J., Dürrwächter, U., Adler, M., and Klosinski, G. (2004). Eye movement control during single-word reading in dyslexics. *J. Vis.* 4, 388–402. doi: 10.1111/j.1755-3768.2009.01523.x
- Marx, C., Hawelka, S., Schuster, S., and Hutzler, F. (2015). An incremental boundary study on parafoveal preprocessing in children reading aloud: parafoveal masks overestimate the preview benefit. *J. Cog. Psychol* 27, 549–561. doi: 10.1080/20445911.2015.1008494
- McConkie, G. W., and Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Percept. Psychophys.* 17, 578–586. doi: 10.3758/BF03203972
- Moll, K., and Landerl, K. (2010). *SLRT-II. Lese- und Rechtschreibtest*. Bern: Huber.
- Norton, E. S., and Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: implications for understanding and treatment of reading disabilities. *Annu. Rev. Psychol.* 63, 427–452. doi: 10.1146/annurev-psych-120710-100431
- Pagán, A., Blythe, H. I., and Liversedge, S. P. (2015). Parafoveal preprocessing of word initial trigrams during reading in adults and children. *J. Exp. Psycho. Learn. Mem. Cogn.* 42, 411–432. doi: 10.1037/xlm0000175



- Pan, J., Yan, M., Laubrock, J., Shu, H., and Kliegl, R. (2013). Eye-voice span during rapid automatized naming of digits and dice in Chinese normal and dyslexic children. *Dev. Sci.* 16, 967–979. doi: 10.1111/desc.12075
- R Core Team (2015). *The R Project for Statistical Computing*. Available at: <http://www.R-project.org/>
- Rau, A. K., Moeller, K., and Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: an eye-tracking study. *Sci. Stud. Read.* 18, 224–233. doi: 10.1080/10888438.2013.857673
- Rau, A. K., Moll, K., Snowling, M. J., and Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *J. Exp. Child Psychol.* 130, 92–105. doi: 10.1016/j.jecp.2014.09.012
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cogn. Psychol.* 7, 65–81. doi: 10.1016/0010-0285(75)90005-5
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *J. Exp. Child Psychol.* 41, 211–236. doi: 10.1016/0022-0965(86)90037-8
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* 62, 1457–1506. doi: 10.1080/17470210902816461
- Reichle, E. D., Pollatsek, A., Fisher, D. L., and Rayner, K. (1998). Towards a model of eye movement control in reading. *Psychol. Rev.* 105, 125–157. doi: 10.1037/0033-295X.105.1.125
- Reichle, E. D., Rayner, K., and Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: comparison to other models. *Brain Behav. Sci.* 26, 445–476. doi: 10.1017/S0140525X03000104
- Schotter, E. R., Angele, B., and Rayner, K. (2012). Parafoveal processing in reading. *Atten. Percept. Psychophys.* 74, 5–35. doi: 10.3758/s13414-011-0219-2
- Seymour, P. H., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Brit. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859
- Share, D. L. (2008). On the Anglocentricities of current reading research and practice: the perils of overreliance on an “outlier” orthography. *Psychol. Bull.* 134, 584–615. doi: 10.1037/0033-2909.134.4.584
- Sperlich, A., Schad, D. J., and Laubrock, J. (2015). When preview information starts to matter: development of the perceptual span in German beginning readers. *J. Cogn. Psychol.* 27, 511–530. doi: 10.1080/20445911.2014.993990
- Tiffin-Richards, S. P., and Schroeder, S. (2015). Children’s and adults’ parafoveal processes in German: phonological and orthographic effects. *J. Cogn. Psychol.* 27, 531–548. doi: 10.1080/20445911.2014.999076
- Venables, W. N., and Ripley, B. D. (2002). *MASS: Modern Applied Statistics with S. Fourth edition*. New York, NY: Springer.
- Wei, T. (2013). *Visualization of a Correlation Matrix (“corrplot”;* R Package, Version 0.73). Available at: <https://github.com/taiyun/corrplot>
- Wickham, H., and Chang, W. (2015). *An Implementation of the Grammar of Graphics. (“ggplot2”;* R Package, Version 1.0.1). Available at: <http://ggplot2.org>
- Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Appl. Psychol.* 14, 1–33. doi: 10.1017/S0142716400010122
- Wolf, M. (1991). Naming speed and reading: the contribution of the cognitive neurosciences. *Read. Res. Q.* 26, 123–141. doi: 10.2307/747978
- Wolf, M., Bowers, P. G., and Biddle, K. (2000). Naming-speed processes, timing, and reading: conceptual review. *J. Learn. Disab.* 33, 387–407. doi: 10.1177/002221940003300409
- 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
- Zoccolotti, P., De Luca, M., Judica, A., and Spinelli, D. (2008). Isolating global and specific factors in developmental dyslexia: a study based on the rate and amount model (RAM). *Exp. Brain Res.* 186, 551–560. doi: 10.1007/s00221-007-1257-9

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Marx, Hutzler, Schuster and Hawelka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Anomalous Cerebellar Anatomy in Chinese Children with Dyslexia

Ying-Hui Yang<sup>1,2†</sup>, Yang Yang<sup>3,4†</sup>, Bao-Guo Chen<sup>5</sup>, Yi-Wei Zhang<sup>6</sup> and Hong-Yan Bi<sup>1\*</sup>

<sup>1</sup> Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China, <sup>2</sup> The University of Chinese Academy of Sciences, Beijing, China, <sup>3</sup> Department of Linguistics, University of Hong Kong, Hong Kong, China, <sup>4</sup> State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong, China, <sup>5</sup> School of Psychology, Beijing Normal University, Beijing, China, <sup>6</sup> School of Labor and Human Resources, Renmin University of China, Beijing, China

## OPEN ACCESS

### Edited by:

Simone Aparecida Capellini,  
São Paulo State University, Brazil

### Reviewed by:

Angela Jocelyn Fawcett,  
Swansea University, UK  
Elisabete Castelon Konkiewitz,  
Universidade Federal da Grande  
Dourados, Brazil

### \*Correspondence:

Hong-Yan Bi  
bihy@psych.ac.cn

<sup>†</sup> These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 11 November 2015

**Accepted:** 21 February 2016

**Published:** 18 March 2016

### Citation:

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

The cerebellar deficit hypothesis for developmental dyslexia claims that cerebellar dysfunction causes the failures in the acquisition of visuomotor skills and automatic reading and writing skills. In people with dyslexia in the alphabetic languages, the abnormal activation and structure of the right or bilateral cerebellar lobes have been identified. Using a typical implicit motor learning task, however, one neuroimaging study demonstrated the left cerebellar dysfunction in Chinese children with dyslexia. In the present study, using voxel-based morphometry, we found decreased gray matter volume in the left cerebellum in Chinese children with dyslexia relative to age-matched controls. The positive correlation between reading performance and regional gray matter volume suggests that the abnormal structure in the left cerebellum is responsible for reading disability in Chinese children with dyslexia.

**Keywords:** developmental dyslexia, Chinese, voxel-based morphometry, left cerebellum, gray matter

## INTRODUCTION

Developmental dyslexia (DD) is characterized by an unexpected difficulty in reading, which is not explained by intellectual impairment, sensory deficits, lack of adequate schooling opportunities, or neurological illness. It is a common developmental disorder affecting 5–18% of school-aged children (Snowling, 2000). The phonological deficit theory is widely accepted in alphabetic languages, suggesting that dyslexics have specific sound manipulation impairments, which affects their auditory memory, word recall, and sound association skills when processing speech (Shaywitz et al., 1998; Ramus, 2001; Hoeft et al., 2007a,b; Ramus and Ahissar, 2012; Boets et al., 2013; Ramus et al., 2013). However, some researchers argue that deficits at the linguistic level may be an external manifestation of DD, and linguistic deficits can be traced back to a more general perceptual deficit, namely, a dysfunction of magnocellular in sensory pathways. The magnocellular deficit theory asserts that the reading problems derive from impaired sensory processing, caused by abnormal visual, auditory or tactile modalities (Stein and Walsh, 1997). Meanwhile, the cerebellar hypothesis for DD indicates that the cerebellar disorder leads to motor control, or automatization difficulties that would subsequently cause reading and writing problems (Nicolson et al., 2001). Furthermore, as the cerebellum receives massive input from various magnocellular systems, it is also proposed that the cerebellar deficit should be unified under the generally magnocellular theory of dyslexia (Stein, 2001). The cerebellar deficit hypothesis has been supported extensively by behavioral, neuroanatomical and neuropsychological studies (Nicolson et al., 1999; Brown et al., 2001; Rae et al., 2002; Stoodley et al., 2008). For example, compared with typically developing readers, children, and adults with dyslexia performed poorer in several tasks related to cerebellar

function, such as time estimation (Wolff et al., 1990), automatic balance (Brookes et al., 2010), speed processing (Nicolson et al., 1995), and implicit motor learning (Stoodley et al., 2008). Using positron emission tomography (PET), Nicolson et al. (1999) directly examined the cerebellar function in learning novel sequences and executing prelearned sequences in adults with dyslexia. It was found that relative to control group, dyslexics exhibited reduced activation in the right cerebellar cortex during both learning novel sequences and executing prelearned sequences. Menghini et al. (2006) found that in alphabetic languages, adults with dyslexia showed increased activation in the right cerebellum than the normal control group. Adults with dyslexia showed less activation in the left cerebellum than the control group during word and pseudoword reading (McCrory et al., 2000). In a meta-analysis, Linkersdörfer et al. (2012) identified that the structural abnormalities were related to functional abnormalities in bilateral cerebellar lobes in dyslexics. Anatomical cerebellar abnormality in dyslexia was supported by the evidence showing the absence of cerebellar asymmetry in dyslexic children (Rae et al., 2002). Structural magnetic resonance imaging (MRI) studies of dyslexia using voxel-based morphometry (VBM) have revealed regional gray matter reductions in the right cerebellum (Brown et al., 2001; Pernet et al., 2009), or in the bilateral cerebellum (Eckert et al., 2003, 2005; Brambati et al., 2004; Kronbichler et al., 2008). However, Richlan et al. (2013) did not find significant cerebellar volume reduction in dyslexics. In general, most of the above studies showed abnormal cerebellar activation and structure in right or lateral lobules in dyslexia of alphabetic languages.

But different from alphabetic languages, Chinese is a logography without clear grapheme-phoneme rules, and Chinese characters are square-shaped with more visually complicated structures. These differences may lead to different neural basis of Chinese character processing (Bolger et al., 2005; Tan et al., 2005; Wu et al., 2012). For example, several neuroimaging studies of normal readers have shown activation in the left middle frontal gyrus (MFG; BA 9) during Chinese reading, which was thought to be specialized for orthography-to-phonology transformation in Chinese processing (Tan et al., 2005; Siok et al., 2008), while in alphabetic-language reading, the left posterior temporal lobe was recruited to perform the conversion of written symbols (letters) into phonological units of speech (phonemes) (Booth et al., 2002). The right parietal and inferior occipital cortices were thought to be engaged in visuospatial analysis in processing Chinese characters (Tan et al., 2001; Chen et al., 2002), while the right superior frontal gyrus, right parietal regions and bilateral cuneus were known to be critical for visuospatial processes in processing alphabetic words (Haxby et al., 1995; Lepage et al., 2000). Thus, there is a discrepancy of the neural basis between Chinese and alphabetic languages processing.

In neuroimaging studies of dyslexia, Siok et al. (2004, 2008) have found that, compared with typically developing reading, dyslexic reading in Chinese was characterized by reduced activation in the left MFG during homophonic (Siok et al., 2004) or rhyming (Siok et al., 2008) judgment. While some studies found that the activation in the left temporoparietal and occipitotemporal regions was abnormally decreased during

alphabetic-language reading in dyslexia (Horwitz et al., 1998; Johansson, 2006; Schlaggar and McCandliss, 2007). Hu et al. (2010) found a similar pattern of brain activity in semantic decision tasks in Chinese and English people with dyslexia, both of them showing reduced activation in the left angular gyrus, left middle frontal cortex, and left occipitotemporal regions relative to normal readers even though Chinese and English normal readers displayed distinct activation in the brain. Therefore, it is not clear whether the neural basis of deficits in DD varies across languages.

As for research in cerebellar deficits of Chinese dyslexia, one behavioral study found that Chinese dyslexia had problems in implicit motor learning when they responded with their left hands, whereas this problem disappeared when using their right hands. In contrast, age-matched children showed significant implicit motor learning when responded with either hand (Yang and Hong-Yan, 2011). The observation of the left-hand response deficits during implicit motor learning in Chinese dyslexia led the researchers to speculate that Chinese dyslexia is likely to be associated with left cerebellar dysfunction, which may be different from the previous studies of cerebellar deficits in alphabetic-language dyslexics (Nicolson et al., 1999; Menghini et al., 2006). Yang et al. (2013) have performed a functional magnetic resonance imaging (fMRI) study to examine cerebellar function in an implicit motor learning task in children with and without dyslexia. The results indicated that Chinese children with dyslexia had significantly higher activity in the left cerebellum compared with age-matched normal children (Yang et al., 2013). Thus, these findings suggested different cerebellar deficits in Chinese and alphabetic dyslexia. As discussed previously (Linkersdörfer et al., 2012), functional deficits usually come along with structural defects. Therefore, this study aimed to determine whether there were structural abnormalities in the left cerebellum in Chinese dyslexia.

## MATERIALS AND METHODS

### Participants

Nine dyslexic children (3 boys, mean age = 12.6 years,  $SD = 0.6$ ) and 14 normal control group (6 boys, mean age = 12.3 years,  $SD = 1.0$ ) took part in the study. The children were recruited from ordinary primary schools in Beijing. Two tests that were widely used for screening Mandarin-speaking Chinese children with dyslexia were adopted: the Raven Standard Progressive Matrices (Raven et al., 1996), and the Character Recognition Test and Assessment Scale (Wang and Tao, 1993). The vocabulary test used in the present study is a standardized vocabulary test for screening DD in Mainland China. In this test, the children were required to write a compound word using a given Chinese character. Each correctly used character was given one point. It includes 210 characters which are divided into 10 sub-groups based on their reading difficulty, which corresponds with the standard difficulty coefficient. The score for each sub-group was calculated by multiplying the total points by the corresponding coefficient of reading difficulty. The final score was measured by adding the total score of 10 sub-groups' and the constant which

was the number of characters almost all children in the same grade could recognize. It is obvious that, raw scores of the test are measured based on the standard difficulty coefficient. And many previous studies with Chinese dyslexia used the raw scores of this vocabulary test to select dyslexia children (Shu et al., 2006; Meng et al., 2007; Wang et al., 2010; Yang and Hong-Yan, 2011; Liu et al., 2012, 2013; Yang et al., 2013; Qian and Bi, 2014; Zhao et al., 2014, 2015). Besides, in this current study, the inclusionary criterion for dyslexics was their written score at least 1.5 standard deviation (*SD*) below the average score of all participants, not a fixed score. In addition, a rapid digit naming task was administered, in which five digits (2, 7, 4, 9, and 6) were presented in random order on a 6 × 5 column grid. All children were asked to read the 30 Arabic digits twice as quickly and accurately as possible. The reading time was recorded. This test was adopted to measure children's rapid automatized naming ability (Zhao et al., 2014). The inclusion criteria for selecting DD were as follows: (1) their reading scores were at least 1.5 *SD* below the average score of age-matched children; and (2) the children had an IQ score higher than 85 in the Raven test. Detailed information of participants was presented in **Table 1**. None of the children suffered from attention deficit/hyperactivity disorder (ADHD) according to the scores of the Chinese Classification of Mental Disorder 3 (CCMD-3). All children had no history of sensory deficits or neurological or psychiatric illness. All participants were right-handed based on the Handedness Inventory (Department of Neurology, Beijing Medical University Hospital). The study was approved by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences, and written informed consent was obtained from all participants' guardians.

## MRI Acquisition

The MRI data were obtained on a 3 Tesla Siemens MAGNETOM Trio scanner (Siemens, Erlangen, Germany) with a standard head coil. A T1-weighted gradient-echo planar imaging (EPI) sequence was used for each subject's high-resolution whole-brain images, with the repetition time = 25 ms, echo time = 30 ms, field of view = 25 mm, matrix size = 256 × 256, voxel size = 1 mm × 1 mm × 1 mm, and 128 non-contiguous (gapped) slices of 4-mm thickness.

## Image Processing and Analysis

Image analysis was performed using SPM8 software<sup>1</sup> (Statistical Parametric Mapping) in MATLAB 7.8 (R2009a) (Math Works, Natick, MA, USA). T1-weighted images were analyzed using

the VBM8 toolbox<sup>2</sup>. Spatial normalization was achieved by registering each image to the standard T1 template implanted in SPM8, based on the Montreal Neurological Institute (MNI) stereotactic space. In the present analysis, the first step in spatially normalizing each image involved matching the image by estimating the optimum parameter affine transformation, and then estimated the coefficients of the basic functions to minimize the residual squared difference between the image and the template by using the non-linear registration. The spatially normalized images were partitioned into gray matter, white matter and cerebral spinal fluid with a resampling at 1 mm × 1 mm × 1 mm resolution, using a modified mixture cluster analysis technique. The segmented images were then modulated (to correct for local expansion or contraction) by dividing with the Jacobian of the warp field. The modulated segmented gray matter images were smoothed with an isotropic Gaussian kernel with a full width at half maximum of 8 mm. The actual volumes of the entire normalized, segmented, and restored segmented images were determined by adding the voxel volumes (1 mm × 1 mm × 1 mm), and multiplying by each voxel value. Intracranial volume was determined by adding the gray matter, white matter, and cerebrospinal fluid space volumes.

No participants from the two groups were excluded because of movement artifacts or incomplete brain scans. We conducted whole-brain gray matter volume analysis using SPM8. Statistical parametric maps of whole-brain VBM analyses were displayed on a template brain. Group differences in the whole-brain gray matter were assessed with the two-sample *t* statistic within SPM software, with total intracranial volume as a covariant; the corrected statistical threshold was set at  $p < 0.001$ , corrected by AlphaSim correction, extent threshold  $k = 111$  voxels. Coordinate points of regions of interest (ROI) were generated and labeled based on regions showing significant group differences in gray matter volume in the above VBM analysis. ROI analysis was used to confirm the statistical parametric map results and to perform correlational analyses by extracting mean gray matter volumes from all participants. Average gray matter volumes of these ROIs for each individual were extracted using the MarsBaR toolbox<sup>3</sup> with 6 mm radius sphere centered at the peak of the group difference.

Structural covariance analysis was used to investigate whether there was significant covariance in gray matter volume among the left cerebellum and other brain regions showing group differences

<sup>2</sup><http://dbm.neuro.uni-jena.de/vbm.html>

<sup>3</sup><http://marsbar.sourceforge.net>

<sup>1</sup>[www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm)

**TABLE 1 | Information concerning the dyslexia and control groups.**

	Dyslexic readers ( <i>n</i> = 9)	Normal readers ( <i>n</i> = 14)	<i>p</i>	$\eta^2$
	Mean ( <i>SD</i> )	Mean ( <i>SD</i> )		
Age (years)	12.6 (0.61)	12.3 (1.0)	>0.05	0.15
Vocabulary (standard score)	2487.39 (149.94)	3101.52 (123.36)	<0.001	0.794
Raven (standard score)	110.35 (15.52)	114.26 (10.35)	>0.05	0.074
Time of digit naming (seconds)	10.63 (1.03)	9.34 (1.01)	<0.01	0.294



in volumes in the VBM analysis. This method has been used previously to examine gray matter correlations between regional volumes of circumscribed brain regions (Mechelli et al., 2005; Pernet et al., 2009; Liu et al., 2013). For structural covariance analysis, the threshold was set at  $p < 0.005$  uncorrected, extent threshold  $k = 40$  voxels. Total gray matter volume, age, and gender were included as covariates in the follow-up analyses for all participants to investigate whether the group differences in the structural co-variation remained significant after regressing out these factors of no interest.

## RESULTS

### Comparison of Gray Matter Between Children With Dyslexia and Controls

Voxel-based morphometry analysis showed that regional gray matter volume in the left cerebellar posterior lobe was significantly smaller in dyslexia than that in controls (MNI coordinates:  $-42/-73/-34$ ; AlphaSim corrected,  $t = 4.61$ , cluster threshold  $p < 0.001$ ,  $k = 111$  voxels; as shown in **Figure 1**). There was no significant difference in total gray matter [ $t = -0.17$ ,  $p = 0.86$ ] and whole brain volume [ $t = 0.81$ ,  $p = 0.42$ ] between children with dyslexia and controls.

We reduced the statistical threshold to  $p < 0.005$  uncorrected and extent threshold  $k = 40$  voxels. VBM analysis revealed decreased gray matter volume in more widespread regions in children with dyslexia compared to controls, including the left superior temporal region (BA 38), left lateral orbitofrontal cortex

(LOFC; BA 47), left MFG (BA 9), left postcentral gyrus (BA 41), and some right brain regions, such as the right cerebellum, right superior frontal gyrus (BA 6), and right fusiform (BA 37). Some regions showed an increase in gray matter in children with dyslexia compared with normal controls located in the right middle temporal gyrus (BA 21), right superior occipital gyrus (BA 18), and right precuneus (BA 7; as shown in **Table 2**).

### Structure-Behavior Correlation and Structural Covariance Analysis

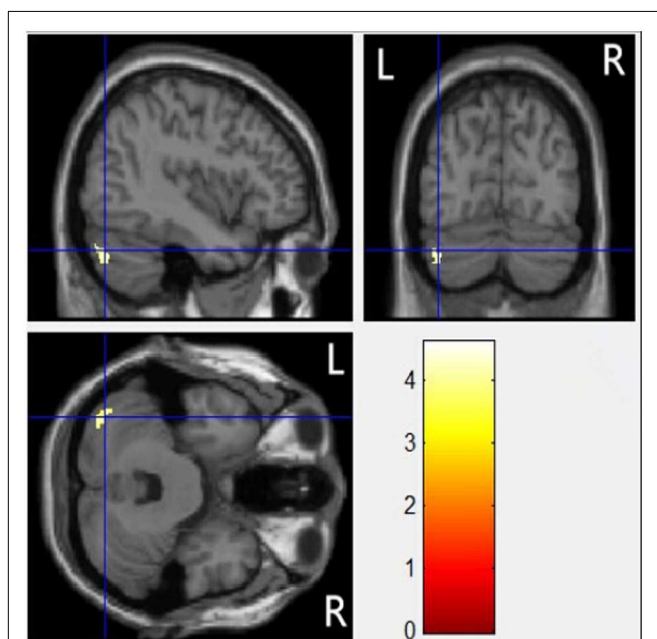
Gray matter volumes in the left cerebellum were correlated with reading scores (vocabulary) for all participants ( $r = 0.62$ ,  $p = 0.002$ ; as shown in **Figure 2**).

Gray matter correlations between the left cerebellum and other regions were also analyzed for all participants. The gray matter of the left cerebellum was significantly correlated with gray matter of several regions in the left hemisphere involving the left superior temporal gyrus (STG;  $r = 0.52$ ,  $p = 0.010$ ), the left LOFC ( $r = 0.48$ ,  $p = 0.020$ ), and the left postcentral cortex ( $r = 0.61$ ,  $p = 0.002$ ) and meanwhile was marginally correlated with the left MFG ( $r = 0.39$ ,  $p = 0.062$ ). Critically, the left cerebellum showed significant positive correlations with brain areas in the right hemisphere involving the right cerebellum ( $r = 0.69$ ,  $p = 0.000$ ) and right fusiform gyrus ( $r = 0.45$ ,  $p = 0.031$ ), and significant negative correlations with the right precuneus ( $r = -0.48$ ,  $p = 0.020$ ). These co-variations were still significant after regressing out the effect of total gray matter, age, and gender. In sum, children with less gray matter in the left cerebellar region tended to show decreased volumes in the right cerebellum, right fusiform, and increased volumes in the right precuneus.

## DISCUSSION

The major finding of the present study was that children with dyslexia displayed a significantly reduced gray matter volume in the left cerebellum, and gray matter volume of the left cerebellum was positively correlated with vocabulary score. Volume of the left cerebellum was positively related to the volume of the right cerebellum, the right fusiform gyrus, and negatively associated with the volume of the right precuneus.

The remarkable finding was decreased gray matter volume of the left cerebellum in Chinese dyslexic children. It was proposed that structural abnormality came along with functional abnormality deficits (Linkersdörfer et al., 2012). Combined with Yang et al. (2013) previous fMRI study which presented that the left cerebellum showed significantly higher activation in Chinese DD during motor sequence learning, we could see that Chinese people with dyslexia indeed have both functional and anatomical deficits in the left cerebellum. What's more, the cerebellum deficits of Chinese reading dyslexia in this study was found in the left side, different from the right cerebellar abnormality found in previous studies of alphabetic-language dyslexics (Nicolson et al., 1999; Brown et al., 2001; Menghini et al., 2006; Pernet et al.,



**FIGURE 1 | Statistical parametric maps of whole-brain VBM analyses displayed on a template brain.** A region in the left posterior cerebellum exhibited significantly reduced gray matter volume in the dyslexic group compared with controls.

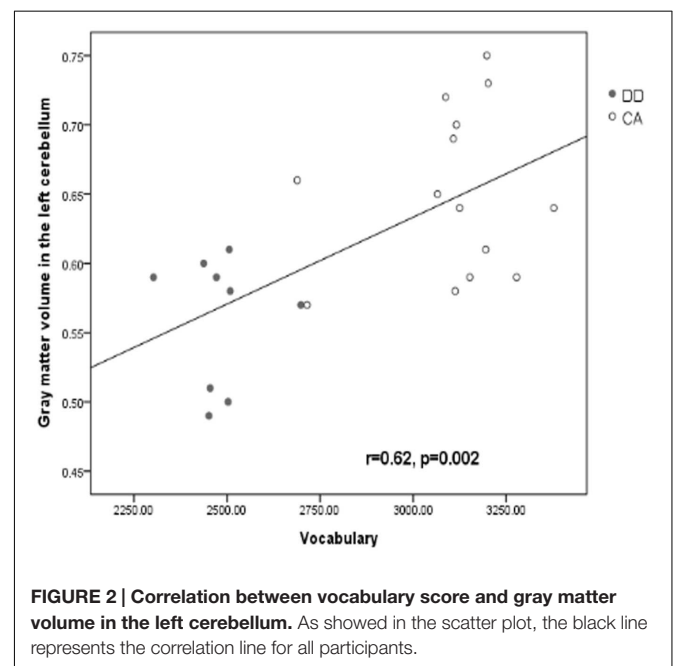
TABLE 2 | Gray matter volume comparisons.

Hemisphere lobe	Anatomical location	MNI coordinate			Z-score	Volume (voxels)
		x	y	z		
Dyslexics < controls						
R(cerebellum)	Cerebellum posterior	45	−63	−36	3.86	139
L(temporal)	Superior temporal gyrus	−40	10	−22	3.67	72
L(frontal)	Lateral orbitofrontal cortex	−49	16	3	3.66	53
L(frontal)	Middle frontal gyrus	−33	49	15	4.19	46
L(parietal)	Postcentral gyrus	−57	−21	14	4.31	42
R(frontal)	Superior frontal gyrus	25	12	60	4.81	52
R(occipital)	Fusiform gyrus	36	−67	−15	5.46	40
Dyslexics > controls						
R(temporal)	Middle temporal gyrus	42	9	−40	7.03	55
R(occipital)	Superior occipital gyrus	18	−87	22	4.18	50
R(parietal)	Precuneus	12	−52	42	6.39	45

2009). This might reflect a unique neural mechanism of Chinese processing.

Much research has shown that Chinese and alphabetic languages have different neural basis and networks. For example, in word-form processing, an additional right middle occipital gyrus was utilized to process the holistic visuospatial configuration of Chinese characters, while this region was not activated during alphabetic-language reading (Bolger et al., 2005; Tan et al., 2005; Cao et al., 2010). Besides, during phonological processing, the left MFG was associated with the phonological processes of character-syllable mapping in Chinese (Siok et al., 2008). While in alphabetic languages, many left brain regions such as the left inferior frontal gyrus (IFG), the left STG, and left inferior parietal lobule (IPL) were involved in phonological processing. The left IFG was thought to be associated with alphabetic word reading (Fiez et al., 1999; Cornelissen et al., 2009) as well as phoneme manipulation and phonological rehearsal before speech production (Fiez, 1997; Price, 2010). The left IPL was responsible for grapheme-phoneme conversion (Paulesu et al., 2000; Booth et al., 2003), and the left STG has been found to be associated with fine-grained phonological representation (Booth et al., 2003; Temple et al., 2003; Nakamura et al., 2006). In the current study, our major finding was that Chinese children with dyslexia exhibited the reduction of gray matter volume in the left cerebellum, while dyslexia of alphabetic languages seemingly usually presented the right cerebellum deficits (Brown et al., 2001; Pernet et al., 2009). So our finding provides a new evidence to demonstrate the different neural basis for different language writing systems.

The present results also showed the volume of left cerebellum was correlated with vocabulary scores, which may suggest the cerebellum is closely related to reading ability, supporting cerebellar deficit hypothesis (Nicolson et al., 2001). It has been reported that the cerebro-cerebellar circuitry regulates the higher order cognitive processing, such as language processing, executive control functions, and working memory (Murdoch, 2010). Clinical and anatomical studies have shown crossed reciprocal connections of the Crus I/II (posterior cerebellar lobe)



with the dorsolateral prefrontal cortex (BA 9/46) (Petrides and Pandya, 1999; Middleton and Strick, 2001), and inferotemporal and posterior parietal cortices (BA 7) (Ramnani, 2006; Jissendi et al., 2008). The perspective on cerebellar involvement in language stems from the cerebro-cerebellar interactions in linguistic functions (Cotterill, 2001; Marien et al., 2001; Schlaggar and McCandliss, 2007). The reciprocal connections between the cerebellum and Broca's language area (lateral temporal and inferior frontal cortices) have also been demonstrated by functional neuroimaging studies, indicating that the cerebellum may be engaged in modulating both language production and comprehension (Petersen et al., 1989; Gebhart et al., 2002; Hubrich-Ungureanu et al., 2002; Jansen et al., 2005; Murdoch and Whelan, 2007; Stoodley and Schmahmann, 2009). It was the left cerebellum volume correlated with reading abilities in this study

in Chinese dyslexia. The left cerebellum should be connected with the right cerebral cortex (Schmahmann, 1996; Middleton and Strick, 2001; Salmi et al., 2010). In fact, numerous studies have shown greater involvement of the right hemisphere in Chinese language processing, such as the right ventral occipital cortex, right superior and right IPLs (Tan et al., 2000; Liu and Perfetti, 2003; Xi et al., 2010; Wu et al., 2012). Therefore, we conjecture the deficits in the left cerebellum of Chinese dyslexic children might affect language processing in Chinese DD via influencing the activation of the right language-related cerebral cortex which is essential for Chinese reading, through the cerebro-cerebellar network.

The results also showed that the volume of the left cerebellum was positively related with the volume of the right fusiform gyrus and negatively correlated with right precuneus. Previous Chinese studies have demonstrated that the right ventral occipital cortex showed greater activation in orthographic processing of Chinese characters (Tan et al., 2001), or visual spatial analysis of characters (Liu and Perfetti, 2003; Bolger et al., 2005). On the other hand, the right precuneus (BA 7) was reported to be involved in visuospatial processes and reinstatement of visual images associated with remembered words (Kjaer et al., 2001; Cavanna and Trimble, 2006). The right precuneus was also proved to be a part of the reciprocal neuroanatomical connections between the cerebellum and cerebral cortex (prefrontal, temporal and parietal cortices) (Buckner et al., 2011; Stoodley, 2012). Chinese character is more complex than alphabetic languages, which needs more visuospatial processing. The correlations between the left cerebellum and right fusiform gyrus and right precuneus might further suggest that the left cerebellum might play a role in the visual processing of Chinese character during reading through its connection with visuospatial processing regions. This finding might also lend some support to the cerebellum hypothesis (Nicolson et al., 2001). An alternative explanation about the negative correlation between the left cerebellum and right precuneus may be resolved by evidence from human (Chang et al., 2005) and animal (Rosen et al., 2000) studies indicating that connectivity disorders between neighboring brain regions may lead to atrophic changes. These changes might include decreased gray matter. The negative correlation of increased gray matter in the right precuneus could be a result of experience-dependent structural remodeling of cortical circuits underlying the acquisition of skills (May, 2011).

What's more, in the current study, some regions also showed gray matter volume reductions in dyslexic children when the statistical threshold reduced to  $p < 0.005$ , such as the left

MFG and the LOFC. Based on the previous findings, the left MFG played a particularly important role in Chinese phonological processing (Tan et al., 2005). And the structure and function of the left MFG of Chinese dyslexics were reported abnormal in an fMRI study (Siok et al., 2008) which was consistent with the present result, suggesting phonological processing deficits in Chinese dyslexia. Besides, the LOFC was thought to be a region involved in spatial attention processing (Armony and Dolan, 2002), and this region was often associated with attention deficit in dyslexia (Facoetti et al., 2000). The current results including structural abnormalities in the MFG and LOFC regions were consistent with previous studies suggesting that dyslexics have phonological (Ho et al., 2002; Siok et al., 2008) and attentional (Facoetti et al., 2000) processing deficits. The cerebellar theory also postulates that the cerebellar disorder affecting speech articulation, which would lead to poor phonological representations and phonological skills in dyslexia (Nicolson et al., 2001). As to the relationship between the regions related to phonology/attention and the left cerebellum, the present study can not resolve this, and further research is needed. There is another limitation of the current study, which is the relatively small sample size, so the current findings should be interpreted with caution.

## CONCLUSION

The present study revealed Chinese children with dyslexia exhibited decreased gray matter volume in the left cerebellum. Moreover, the regional gray matter was significantly correlated with reading scores, suggesting the abnormal structure in the left cerebellum is highly associated with reading disability in Chinese children with dyslexia, supporting cerebellum hypothesis.

## AUTHOR CONTRIBUTIONS

YY designed and performed the experiments; YY and Y-WZ collected the data; Y-HY performed data analysis and wrote the manuscript. B-GC and H-YB edited the manuscript.

## ACKNOWLEDGMENT

This research was supported by the grant from Chinese Natural Science Foundation to H-YB (31371044).

## REFERENCES

- Armony, J. L., and Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia* 40, 817–826. doi: 10.1016/S0028-3932(01)00178-6
- 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
- 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
- Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2002). Functional anatomy of intra- and cross-modal lexical tasks. *Neuroimage* 16, 7–22. doi: 10.1006/nimg.2002.1081

- Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. R., and Mesulam, M. M. (2003). The relation between brain activation and lexical performance. *Hum. Brain Mapp.* 19, 155–169. doi: 10.1002/hbm.10111
- 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* 63, 742–745. doi: 10.1212/01.WNL.0000134673.95020.EE
- Brookes, R. L., Tinkler, S., Nicolson, R. I., and Fawcett, A. J. (2010). Striking the right balance: motor difficulties in children and adults with dyslexia. *Dyslexia* 16, 358–373. doi: 10.1002/dys.420
- Brown, W., Eliez, S., Menon, V., Rumsey, J., White, C., and Reiss, A. (2001). Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology* 56, 781–783. doi: 10.1212/WNL.56.6.781
- Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., and Yeo, B. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 2322–2345. doi: 10.1152/jn.00339.2011
- Cao, F., Lee, R., Shu, H., Yang, Y., Xu, G., Li, K., et al. (2010). Cultural constraints on brain development: evidence from a developmental study of visual word processing in Mandarin Chinese. *Cereb. Cortex* 20, 1223–1233. doi: 10.1093/cercor/bhp186
- Cavanna, A. E., and Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. doi: 10.1093/brain/awl004
- Chang, B. S., Ly, J., Appignani, B., Bodell, A., Apse, K. A., Ravenscroft, R. S., et al. (2005). Reading impairment in the neuronal migration disorder of periventricular nodular heterotopia. *Neurology* 64, 799–803. doi: 10.1212/01.WNL.0000152874.57180.AF
- Chen, Y., Fu, S., Iversen, S. D., Smith, S. M., and Matthews, P. M. (2002). Testing for dual brain processing routes in reading: a direct contrast of Chinese character and pinyin reading using fMRI. *J. Cogn. Neurosci.* 14, 1088–1098. doi: 10.1162/089892902320474535
- Cornelissen, P. L., Kringelbach, M. L., Ellis, A. W., Whitney, C., Holliday, I. E., and Hansen, P. C. (2009). Activation of the left inferior frontal gyrus in the first 200 ms of reading: evidence from magnetoencephalography (MEG). *PLoS ONE* 4, 5359. doi: 10.1371/journal.pone.0005359
- Cotterill, R. M. (2001). Cooperation of the basal ganglia, cerebellum, sensory cerebrum and hippocampus: possible implications for cognition, consciousness, intelligence and creativity. *Progr. Neurobiol.* 64, 1–33. doi: 10.1016/S0301-0082(00)00058-7
- Eckert, M. A., Leonard, C. M., Richards, T. L., Aylward, E. H., Thomson, J., and Berninger, V. W. (2003). Anatomical correlates of dyslexia: frontal and cerebellar findings. *Brain* 126, 482–494. doi: 10.1093/brain/awg026
- 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
- Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., and Mascetti, G. G. (2000). Visual-spatial attention in developmental dyslexia. *Cortex* 36, 109–123. doi: 10.1016/S0010-9452(08)70840-2
- Fiez, J. A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex. *Hum. Brain Mapp.* 5, 79–83. doi: 10.1002/(SICI)1097-0193(1997)5:2<79::AID-HBM1>3.0.CO;2-J
- Fiez, J. A., Balota, D. A., Raichle, M. E., and Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* 24, 205–218. doi: 10.1016/S0896-6273(00)80833-8
- Gebhart, A. L., Petersen, S. E., and Thach, W. T. (2002). Role of the posterolateral cerebellum in language. *Ann. N. Y. Acad. Sci.* 978, 318–333. doi: 10.1111/j.1749-6632.2002.tb07577.x
- Haxby, J. V., Ungerleider, L. G., Horwitz, B., Rapoport, S. I., and Grady, C. L. (1995). Hemispheric differences in neural systems for face working memory: a PET-rCBF study. *Hum. Brain Mapp.* 3, 68–82. doi: 10.1002/hbm.460030204
- 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
- Hoef, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., and Martindale, J. L. (2007a). 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
- Hoef, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., et al. (2007b). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behav. Neurosci.* 121, 602. doi: 10.1037/0735-7044.121.3.602
- 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
- Hubrich-Ungureanu, P., Kaemmerer, N., Henn, F. A., and Braus, D. F. (2002). Lateralized organization of the cerebellum in a silent verbal fluency task: a functional magnetic resonance imaging study in healthy volunteers. *Neurosci. Lett.* 319, 91–94. doi: 10.1016/S0304-3940(01)02566-6
- Jansen, A., Flöel, A., Van Randenborgh, J., Konrad, C., Rotte, M., Förster, A. F., et al. (2005). Crossed cerebro-cerebellar language dominance. *Hum. Brain Mapp.* 24, 165–172. doi: 10.1002/hbm.20077
- Jissendi, P., Baudry, S., and Baleriaux, D. (2008). Diffusion tensor imaging (DTI) and tractography of the cerebellar projections to prefrontal and posterior parietal cortices: a study at 3T. *J. Neuroradiol.* 35, 42–50. doi: 10.1016/j.neurad.2007.11.001
- Johansson, B. B. (2006). Cultural and linguistic influence on brain organization for language and possible consequences for dyslexia: a review. *Ann. Dyslexia* 56, 13–50. doi: 10.1007/s11881-006-0002-6
- Kjaer, T. W., Nowak, M., Kjaer, K. W., Lou, A. R., and Lou, H. C. (2001). Precuneus-prefrontal activity during awareness of visual verbal stimuli. *Conscious. Cogn.* 10, 356. doi: 10.1006/ccog.2001.0509
- 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
- Lepage, M., Ghaffar, O., Nyberg, L., and Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. *Proc. Natl. Acad. Sci. U.S.A.* 97, 506–511. doi: 10.1073/pnas.97.1.506
- Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., and Fiebach, C. J. (2012). Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: an ALE meta-analysis. *PLoS ONE* 7, e43122. doi: 10.1371/journal.pone.0043122
- 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
- Liu, Y., and Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: an ERP study of Chinese bilinguals. *Hum. Brain Mapp.* 18, 167. doi: 10.1002/hbm.10090
- Mariën, P., Engelborghs, S., Fabbro, F., and De Deyn, P. P. (2001). The lateralized linguistic cerebellum: a review and a new hypothesis. *Brain Lang.* 79, 580–600. doi: 10.1006/brln.2001.2569
- May, A. (2011). Experience-dependent structural plasticity in the adult human brain. *Trends Cogn. Sci.* 15, 475–482. doi: 10.1016/j.tics.2011.08.002
- McCrory, E., Frith, U., Brunswick, N., and Price, C. (2000). Abnormal functional activation during a simple word repetition task: a PET study of adult dyslexics. *J. Cogn. Neurosci.* 12, 753–762. doi: 10.1162/089892900562570
- Mechelli, A., Friston, K. J., Frackowiak, R. S., and Price, C. J. (2005). Structural covariance in the human cortex. *J. Neurosci.* 25, 8303–8310. doi: 10.1523/JNEUROSCI.0357-05.2005
- Meng, X., Tian, X., Jian, J., and Zhou, X. (2007). Orthographic and phonological processing in Chinese dyslexic children: an ERP study on sentence reading. *Brain Res.* 1179, 119–130. doi: 10.1016/j.brainres.2007.08.046
- Menghini, D., Hagberg, G. E., Caltagirone, C., Petrosini, L., and Vicari, S. (2006). Implicit learning deficits in dyslexic adults: an fMRI study. *Neuroimage* 33, 1218–1226. doi: 10.1016/j.neuroimage.2006.08.024



- Middleton, F. A., and Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *J. Neurosci.* 21, 700–712.
- Murdoch, B. E. (2010). The cerebellum and language: historical perspective and review. *Cortex* 46, 858–868. doi: 10.1016/j.cortex.2009.07.018
- Murdoch, B. E., and Whelan, B. M. (2007). Language disorders subsequent to left cerebellar lesions: a case for bilateral cerebellar involvement in language? *Folia Phoniatr. Logopaed.* 59, 184–189. doi: 10.1159/000102930
- Nakamura, K., Hara, N., Kouider, S., Takayama, Y., Hanajima, R., Sakai, K., et al. (2006). Task-guided selection of the dual neural pathways for reading. *Neuron* 52, 557–564. doi: 10.1016/j.neuron.2006.09.030
- Nicolson, R. I., Fawcett, A. J., Berry, E. L., Jenkins, I. H., Dean, P., and Brooks, D. J. (1999). Association of abnormal cerebellar activation with motor learning difficulties in dyslexic adults. *Lancet* 353, 1662–1667. doi: 10.1016/S0140-6736(98)09165-X
- Nicolson, R. I., Fawcett, A. J., and Dean, P. (1995). Time estimation deficits in developmental dyslexia: evidence of cerebellar involvement. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 259, 43–47. doi: 10.1098/rspb.1995.0007
- 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
- Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., et al. (2000). A cultural effect on brain function. *Nat. Neurosci.* 3, 91–96. doi: 10.1038/71163
- Pernet, C. R., Poline, J. B., Demonet, J. F., and Rousselet, G. A. (2009). Brain classification reveals the right cerebellum as the best biomarker of dyslexia. *BMC Neurosci.* 10:67. doi: 10.1186/1471-2202-10-67
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. (1989). Positron emission tomographic studies of the processing of single words. *J. Cogn. Neurosci.* 1, 153–170. doi: 10.1162/jocn.1989.1.2.153
- Petrides, M., and Pandya, D. N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur. J. Neurosci.* 11, 1011–1036. doi: 10.1046/j.1460-9568.1999.00518.x
- Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N. Y. Acad. Sci.* 1191, 62–88. doi: 10.1111/j.1749-6632.2010.05444.x
- 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
- Rae, C., Harasty, J. A., Dzendrowsky, T. E., Talcott, J. B., Simpson, J. M., Blamire, A. M., et al. (2002). Cerebellar morphology in developmental dyslexia. *Neuropsychologia* 40, 1285–1292. doi: 10.1016/S0028-3932(01)00216-0
- Ramrani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nat. Rev. Neurosci.* 7, 511–522. doi: 10.1038/nrn1953
- Ramus, F. (2001). Outstanding questions about phonological processing in dyslexia. *Dyslexia* 7, 197–216. doi: 10.1002/dys.205
- Ramus, F., and Ahissar, M. (2012). Developmental dyslexia: the difficulties of interpreting poor performance, and the importance of normal performance. *Cogn. Neuropsychol.* 29, 104–122. doi: 10.1080/02643294.2012.677420
- Ramus, F., Marshall, C. R., Rosen, S., and van der Lely, H. K. (2013). Phonological deficits in specific language impairment and developmental dyslexia: towards a multidimensional model. *Brain* 136, 630–645. doi: 10.1093/brain/aww356
- Raven, J. C., Court, J. H., and Raven, J. (1996). *Standard Progressive Matrices*. Oxford: Oxford Psychologists Press.
- 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
- Rosen, G. D., Burstein, D., and Galaburda, A. M. (2000). Changes in efferent and afferent connectivity in rats with induced cerebrotectomicrogyria. *J. Comparat. Neurol.* 418, 423–440. doi: 10.1002/(SICI)1096-9861
- Salmi, J., Pallesen, K. J., Neuvonen, T., Brattico, E., Korvenoja, A., Salonen, O., et al. (2010). Cognitive and motor loops of the human cerebro-cerebellar system. *J. Cogn. Neurosci.* 22, 2663–2676. doi: 10.1162/jocn.2009.21382
- Schlaggar, B. L., and McCandliss, B. D. (2007). Development of neural systems for reading. *Annu. Rev. Neurosci.* 30, 475–503. doi: 10.1146/annurev.neuro.28.061604.135645
- Schmahmann, J. D. (1996). From movement to thought: anatomic substrates of the cerebellar contribution to cognitive processing. *Hum. Brain Mapp.* 4, 174–198. doi: 10.1002/(SICI)1097-0193(1996)4
- 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. doi: 10.1073/pnas.95.5.2636
- 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. doi: 10.1037/0022-0663.98.1.122
- 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
- Snowling, M. J. (2000). *Dyslexia*. Hoboken, NJ: Blackwell Publishing.
- Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia* 7, 12–36. doi: 10.1002/dys.186
- Stein, J., and Walsh, V. (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
- Stoodley, C. J. (2012). The cerebellum and cognition: evidence from functional imaging studies. *Cerebellum* 11, 352–365. doi: 10.1007/s12311-011-0260-7
- Stoodley, C. J., Ray, N. J., Jack, A., and Stein, J. F. (2008). Implicit learning in control. Dyslexic garden-variety poor readers. *Ann. N. Y. Acad. Sci.* 1145, 173–183. doi: 10.1196/annals.1416.003
- Stoodley, C. J., and Schmahmann, J. D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage* 44, 489–501. doi: 10.1016/j.neuroimage.2008.08.039
- Tan, L. H., Laird, A. R., Li, K., and Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. *Hum. Brain Mapp.* 25, 83–91. doi: 10.1002/hbm.20134
- Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., and Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *Neuroimage* 13, 836–846. doi: 10.1006/nimg.2001.0749
- Tan, L. H., Spinks, J. A., Gao, J. H., Liu, H. L., Perfetti, C. A., Xiong, J., et al. (2000). Brain activation in the processing of Chinese characters and words: a functional MRI study. *Hum. Brain Mapp.* 10, 16–27. doi: 10.1002/(SICI)1097-0193
- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2860–2865. doi: 10.1073/pnas.0030098100
- 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.
- Wolff, P. H., Michel, G. F., Ovrut, M., and Drake, C. (1990). Rate and timing precision of motor coordination in developmental dyslexia. *Dev. Psychol.* 26:349. doi: 10.1037/0012-1649.26.3.349
- Wu, C. Y., Ho, M. H. R., and Chen, S. H. A. (2012). A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. *Neuroimage* 63, 381–391. doi: 10.1016/j.neuroimage.2012.06.047
- Xi, J., Zhang, L., Shu, H., Zhang, Y., and Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience* 170, 223–231. doi: 10.1016/j.neuroscience.2010.06.077
- 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., and Hong-Yan, B. (2011). Unilateral implicit motor learning deficit in developmental dyslexia. *Int. J. Psychol.* 46, 1–8. doi: 10.1080/00207594.2010.509800
- Zhao, J., Qian, Y., Bi, H. Y., and Coltheart, M. (2014). The visual magnocellular-dorsal dysfunction in Chinese children with developmental dyslexia impedes Chinese character recognition. *Sci. Rep.* 4:7068. doi: 10.1038/srep07068
- Zhao, J., Yang, Y., Song, Y. W., and Bi, H. Y. (2015). Verbal short-term memory deficits in Chinese children with dyslexia may not be a problem with the activation of phonological representations. *Dyslexia* 21, 304–322. doi: 10.1002/dys.1516

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Yang, Yang, Chen, Zhang and Bi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Role of Reading Fluency in Children's Text Comprehension

Marta Álvarez-Cañizo\*, Paz Suárez-Coalla and Fernando Cuetos

Department of Psychology, University of Oviedo, Asturias, Spain

## OPEN ACCESS

### Edited by:

Simone Aparecida Capellini,  
São Paulo State University, Brazil

### Reviewed by:

Christelle Declercq,  
Université de Reims  
Champagne-Ardenne, France  
Vera Lúcia Orlandi Cunha,  
São Paulo State University, Brazil

### \*Correspondence:

Marta Álvarez-Cañizo  
alvarezcanmarta@uniovi.es

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 28 July 2015

**Accepted:** 09 November 2015

**Published:** 27 November 2015

### Citation:

Álvarez-Cañizo M, Suárez-Coalla P  
and Cuetos F (2015) The Role  
of Reading Fluency in Children's Text  
Comprehension.  
Front. Psychol. 6:1810.  
doi: 10.3389/fpsyg.2015.01810

Understanding a written text requires some higher cognitive abilities that not all children have. Some children have these abilities, since they understand oral texts; however, they have difficulties with written texts, probably due to problems in reading fluency. The aim of this study was to determine which aspects of reading fluency are related to reading comprehension. Four expository texts, two written and two read by the evaluator, were presented to a sample of 103 primary school children (third and sixth grade). Each text was followed by four comprehension questions. From this sample we selected two groups of participants in each grade, 10 with good results in comprehension of oral and written texts, and 10 with good results in oral and poor in written comprehension. These 40 subjects were asked to read aloud a new text while they were recorded. Using Praat software some prosodic parameters were measured, such as pausing and reading rate (number and duration of the pauses and utterances), pitch and intensity changes and duration in declarative, exclamatory, and interrogative sentences and also errors and duration in words by frequency and stress. We compared the results of both groups with ANOVAs. The results showed that children with less reading comprehension made more inappropriate pauses and also intersentential pauses before comma than the other group and made more mistakes in content words; significant differences were also found in the final declination of pitch in declarative sentences and in the F0 range in interrogative ones. These results confirm that reading comprehension problems in children are related to a lack in the development of a good reading fluency.

**Keywords:** prosody, text reading, reading comprehension, Spanish, children

## INTRODUCTION

The difficulty understanding written texts is a major cause of school failure because it requires some cognitive abilities, such as previous knowledge activation, inference performance, mental models building, etc., which not all children have. But some children, despite having these skills, fail in understanding texts. Following the Simple View of Reading (Hoover and Gough, 1990), linguistic comprehension and word recognition are needed to achieve reading comprehension. Besides, fluency could facilitate reading comprehension because it frees resources for understanding (Adlof et al., 2006). Therefore, there could be several causes of this poor comprehension; one of them could be that they have not developed a good reading fluency nor have poor decoding skills. Fluent reading involves accuracy, speed and good expression (National Institute of Child Health, and Human Development, 2000). These three characteristics depend on several cognitive processes and are usually achieved in that order, although overlapping. There are some evidences about the relationship between text reading fluency and reading comprehension,

Kim and Wagner (2015) showed that the role of text reading fluency walks together the reading comprehension improvement.

Text reading accuracy is one of the more decisive factors in reading comprehension. Thus, if a child makes many mistakes he cannot understand what he is reading. Moreover, there are some words that are more difficult to read, such as long words (Muncer et al., 2014), low frequency words (Fischer-Baum et al., 2014), words with few orthographic neighbors (Laxon et al., 2002), late age of acquisition words (Cuetos and Barbón, 2006; Monaghan and Ellis, 2010; Davies et al., 2014) or words with complex syllabic structure (Taft, 1979; Rouibah et al., 2000). These kinds of words are often read with less accuracy, and that could affect comprehension.

Speed is also an important part of the reading process. Perfetti (1985) in his Verbal Efficiency Theory states that readers who lack efficient word identification procedures are at risk for comprehension failure. If readers are quick and accurate in identifying words, they will have more attentional resources to devote to understanding what they are reading. Therefore, slowness is also an additional problem, as it consumes working memory and, thus, prevents the reader from thinking about the text while reading. Consequently, slow reading especially affects long sentences, because when the reader finishes with the last words of the sentence, he has already forgotten the first ones.

Another important process of reading fluency is expressiveness, or prosody. Some authors defined fluency as the ability to project natural pitch, stress and juncture of spoken words or written text automatically and at a natural rate (Richards, 2000), considering equal prosody and fluency. Besides, other authors consider that fluency is related, not only with appropriate prosody, but with a deep reading understanding (Rasinski, 2004; Ravid and Mashraki, 2007; Hudson et al., unpublished manuscript), prosody becomes a link between fluency and comprehension (Kuhn and Stahl, 2003). However, the direction of the relationship between prosody and comprehension is not clear.

There are some prosodic markers that are indicative of the reader's ability (Dowhower, 1991), such as pausal intrusions, final lengthening in sentences, terminal intonation contours, or stress. Schwanenflugel et al. (2004) purposed five prosodic features: the duration and the variation of appropriate and inappropriate pauses, the pitch sentence and the final declination of pitch in sentence. Good readers usually made fewer and shorter pauses within and between sentences, while less skilled children paused often (Schwanenflugel et al., 2004; Miller and Schwanenflugel, 2006; Benjamin and Schwanenflugel, 2010). Similar results have also been found in studies with adults (Binder et al., 2013), since those with low literacy skills made more word and sentence intrusions compared to the skilled adult readers. Thus, these readers made a higher number of inappropriate pauses while reading and for longer durations.

Moreover, Clay and Imlach (1971), from their study with 7-years-old children, suggested that good readers made not only fewer and shorter pauses, but also had a specific contour pitch in declarative sentences when reading. Similar results were found by Miller and Schwanenflugel (2006, 2008), as they reported that

adequate pitches and better abilities to decode are related. In addition, children who used larger pitch changes and larger end-sentence declinations in reading performed better on reading comprehension than children who used these prosodic features to a lesser extent (Benjamin and Schwanenflugel, 2010).

In transparent orthographic systems, like the Spanish language, children soon get a high level of accuracy, because it is easier to automate the conversion of graphemes into phonemes; this allows children, after the first year of reading learning, a reading accuracy of 95% of words, contrary to opaque languages where the accuracy is about 35% of the words read (Seymour et al., 2003). However, it is possible that early accuracy leads to neglectful reading, and consequently children take a long time to acquire reading fluency. Besides, the prosody is less worked in schools, maybe because of the difficulty of quantifying.

There are scales to measure some specific features of prosody, such as the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Kaminski and Good, 1998), which was validated for assessing reading fluency in DIBELS ORF (DIBELS Oral Reading Fluency; Kame'enui et al., 2006). This scale measures speed, accuracy and pauses when children read a text for 1 min. It is usually used as a measure of the progress of students, who may be at risk for difficulties in future reading comprehension in the educational field. Petscher and Kim (2011) used DIBELS ORF to assess children from different grades; the results did not validate the use of oral reading fluency as the sole measuring of children's reading. Another scale, the Multidimensional Fluency Scale (Rasinski et al., 2009), consists of three subscales to assess phrasing and expression, accuracy and smoothness and pacing. Finally, another scale (Klauda and Guthrie, 2008) assesses several prosodic dimensions, such as expressiveness, phrasing, pace or smoothness. In Spanish, González-Trujillo et al. (2014) created the Reading Fluency Scale in Spanish, based on the Multidimensional Fluency Scale, which consists of the assessment of speed, accuracy and several prosodic features (i.e., volume, intonation, pauses, phrasing, and the reading quality). Children from different grades were assessed using this scale (Calet et al., 2015) and their results showed that also in Spanish, prosodic reading predicts reading comprehension, but depending on the scholar's grade. These scales are very useful in the educational field, but they have some subjectivity.

Today, thanks to programs like Praat (Boersma and Weenink, 2015), it is possible to measure the components of prosody by analyzing the acoustic wave. This is an objective measure of the prosodic features. This software is a tool for phonetic analysis of speech to analyze prosodic aspects such as frequency, intensity or duration.

The aim of this study was to determine which aspects of reading fluency are related to understanding, that is we are interested in the mechanic aspects of reading that could be related to reading comprehension. In this way, several prosodic features, such as pitch, intensity, pauses, duration of syllables and utterances, were collected using Praat software. Besides, words with different lexical frequency and stress were included. A group of children from third and sixth grade with low written comprehension was compared with a group of children with good written comprehension to deal with the objective.



## MATERIALS AND METHODS

### Participants

A total of 103 primary school children (58 females) participated in this study. Forty-six were attending third grade ( $M_{age} = 8.86$ ,  $SD = 0.39$ ) and fifty-seven were sixth grade students ( $M_{age} = 11.89$ ,  $SD = 0.27$ ) in a monolingual school which served children from early childhood (3 years) to high school (17 years). They all had Spanish as their first language and the school served a broadly typical catchment area with the majority children coming from mid-income backgrounds. None of them had developmental, behavioral, or cognitive problems and they also attend school regularly.

This group of children received four expositive texts from the PROLEC-R test (Cuetos et al., 2007), firstly two presented in an oral way (i.e., “El ratel” [“Honey badger”] and “Los vikings” [“Vikings”]), and secondly two in a written format (i.e., “Los indios apaches” [“Apache Indians”] and “Los okapis” [“Okapis”]). The texts were presented in the same order for all the children. Each text was followed by four questions (two inferential and two literals) in order to measure their oral and written comprehension (see **Table 1** for the main means). We selected the children with better results in oral comprehension, with scores between 6 and 8 points out of eight in the oral comprehension texts, in order to ensure that they had the necessary cognitive abilities needed for comprehension. These children were divided into two groups according to their level of reading comprehension, high reading comprehension group when written comprehension results were similar to the above, and low reading comprehension group when they were about three or four points out of eight. In this way two groups of 10 participants in each grade were selected (see **Table 2**). After this selection we had 40 children: 20 children (13 females) with good oral and written comprehension (“Good comprehension”) and 20 (12 females) with only good oral comprehension (“Poor comprehension”). Therefore, there were no significant differences between groups in oral comprehension scores [ $t(19) = 0.79$ ,  $p = 0.48$ ], while there were in written

comprehension scores [ $t(19) = 10.18$ ,  $p < 0.001$ ]. These two final groups were considered the experimental groups, which participated in the second part of the study (as described below). Selected children were also assessed with the reading of words and pseudowords subtests of the PROLEC-R test in order to ensure that everyone had an adequate reading level by age and scholar grade. The poor comprehension group had lower scores than the good comprehension one, but the differences were not significant in both reading of words [ $t(19) = 0.78$ ,  $p = 0.45$ ] and reading of pseudowords [ $t(19) = 2.07$ ,  $p = 0.052$ ] subtests. However, there are significant differences between both groups of third grade in reading of pseudowords subtest [ $t(9) = 2.9$ ,  $p = 0.02$ ], while there are not in sixth grade [ $t(9) = 0.6$ ,  $p = 0.54$ ]. Regarding reading words subtest, there were not significant differences in third grade [ $t(9) = 0.37$ ,  $p = 0.72$ ] nor sixth grade [ $t(9) = 0.73$ ,  $p = 0.49$ ]. PROLEC-R is a standardized battery for the assessment of reading in Spanish children between 6 and 12 years. The reading of words subtest consists of a list of 40 real words, with two or three syllables. In the reading pseudowords subtest children have to read a list of 40 pseudowords, paired by number of syllables, syllabic structure and initial letters with the words list. Children have to read the words and pseudowords aloud; the measurements taken are the number of errors and the time they spent reading each list.

This research was approved by the Ethics Committee of the Psychology Department of the University of Oviedo. Before starting the experimental tasks, the children’s parents received pertinent information about the purpose of the study, the tasks and their duration. Then, written informed consent was received from the parents of participants.

### Material

A narrative text composed by 306 words, titled “El Gigante Egoísta [The Selfish Giant]” (an adaptation of the story by Oscar Wilde), was used. The text was created including declarative (i.e., “Todos eran amigos de Pablo” [“All of them are Pablo’s friends”], “Una mañana el Gigante oyó el trino de un pájaro” [“One morning the Giant heard a bird’s warble”]), exclamatory (i.e., “¡Qué feliz soy aquí!” [“How happy I am here!”], “¡Por fin ha llegado la primavera!” [“Spring has come at last!”]) and interrogative sentences (i.e., “¿Por qué tarda tanto en llegar la primavera?” [“Why does it take so long to get spring?”] and “¿Qué está pasando en mi jardín?” [“What is happening in my garden?”]). It also included eight low frequency words ( $M_{lexicalfrequency} = 13.5$ ; e.g., magnolia [magnolia], secuoya

**TABLE 1 | Results in comprehension questions of the initial participants.**

	Third grade <i>M (SD)</i>	Sixth grade <i>M (SD)</i>
Oral comprehension accuracy (out of eight)	4.3 (1.79)	5.3 (1.66)
Written comprehension accuracy (out of eight)	3.9 (1.78)	5.9 (1.53)

**TABLE 2 | Results in PROLEC-R test and comprehension questions of texts.**

	Third grade		Sixth grade	
	Good comprehension <i>M (SD)</i>	Poor comprehension <i>M (SD)</i>	Good comprehension <i>M (SD)</i>	Poor comprehension <i>M (SD)</i>
Oral comprehension accuracy (out of eight)	5.9 (0.9)	6 (0.7)	6.6 (0.7)	6.8 (0.4)
Written comprehension accuracy (out of eight)	6.1 (0.7)	3 (1.1)	6.9 (0.7)	4.8 (0.9)
Words accuracy (out of 40)	38.9 (0.9)	38.6 (2.1)	39.5 (1.1)	39 (1.6)
Pseudowords accuracy (out of 40)	32.8 (3.5)	29.4 (3.5)	35.2 (2.9)	34 (3.6)

[secoya], subyugado [charmed]), half of them repeated twice, once at the beginning and once at the middle of the text; besides we incorporated 10 words stressed in the penultimate syllable (e.g., palomas [doves], hormigas [ants], tamarindos [tamarinds], estorninos [starlings]), and 10 words stressed in the antepenultimate syllable (e.g., bárbaro [barbarous], mágico [magical], ánfora [amphora], pelícanos [pelicans]), half with low ( $M_{\text{lexical frequency}} = 5.9$ ) and half with high lexical frequency ( $M_{\text{lexical frequency}} = 143.9$ ). The lexical frequency was obtained from the database of Martínez and García (2004), who acquired their frequency from a sample of children's books.

The text was presented on a piece of paper (Times New Roman, 12 point font, double spaced) and the participants had to read it aloud individually in a quiet room. The reading was recorded by an H4n voice recorder and an Ht2-P Audix headset dynamic microphone. Audio recordings were processed offline using Praat software.

## General Assessments

From the .wav files recorded we collected several prosodic parameters using Praat software. First, we analyzed some characteristics of the whole text, and then we extracted six sentences, two declarative, two exclamatory and two interrogative sentences, in order to evaluate different parameters. Finally, we selected eight low frequency words, half of them repeated twice in the text, and eight words with different stresses (on the penultimate and on the antepenultimate syllable) and frequency (high and low).

From the whole text we considered the number of reading mistakes in the content and function words, and the number and duration of intersentential pauses (before commas and full stops) and inappropriate pauses (pauses made in not corresponding places). Also the total pause duration and the total pronunciation time (reading time between pauses) were collected.

Secondly, from the target sentences several measures were used:

Fundamental frequency (F0) measures:

- Range of the first peak (Hz): distance between minimum F0 at the beginning of the sentence and the first peak of F0.
- Total range (Hz): distance between the minimum and the maximum F0 of the sentence.
- Pitch change between different points:
  - From the beginning of the sentence and the first peak (Hz).
  - From the first peak to the end of the sentence (Hz).
  - From the last peak to the end of the sentence (Hz).
  - Between the last syllable and the previous (Hz).

- Slope (Hz/s): declination of the F0 from the first peak to the end of the sentence by time.

Duration measure:

- Phrase-final lengthening (ms): duration of the last syllable of the sentence in comparison with the previous.

Intensity measure:

- Intensity change at the end of the sentence (dB): comparison between the intensity of the last syllable with the previous.

With regards to the target words, the number of errors and mean time of reading were measured. In the case of the words with different stresses, we classified the errors in misreading words and changes in the stress place.

## RESULTS

We compared the results of the different parameters of both groups and grades with ANOVAs using SPSS software. Therefore, we used the results of the measurements described above as dependent variables and the grade (third vs. sixth) and group (poor vs. good comprehension) as the independent variables. We named "Poor comprehension group" as the children with better oral than reading comprehension, and "Good comprehension group" as the children with similar oral and reading comprehension. We made ANOVAs with each dependent variable in order to check which the significant effects were. Only those significant are presented here to facilitate the understanding of the results.

## Text Analyses

We found significant differences by group in the number of reading errors made in content words [ $F(1,36) = 7.85, p = 0.008$ ] and in the number of inappropriate pauses [ $F(1,3) = 4.18, p = 0.048$ ]. Also found was an interaction between group and grade in the number of intersentential pauses before commas [ $F(1,36) = 4.1, p = 0.045$ ]. We performed *post hoc* comparisons using Tukey HSD test and we found that the significant differences in these triple interaction were between third and sixth grades within the poor comprehension group ( $p = 0.024$ ). See Table 3 with the means and SD of these significant effects.

## Sentences Analyses

We discarded for the analysis all the sentences read incorrectly, which was usually regressions in the reading, around 9%. We

**TABLE 3 | Mean and SD of the principal significant effects in the text analyses.**

Significant effect		Poor comprehension <i>M (SD)</i>	Good comprehension <i>M (SD)</i>
Number of reading errors in content words		9.9 (0.65)	7.5 (0.65)
Number of inappropriate pauses		70.5 (4.3)	58 (4.3)
Number of intersentential pauses before comma	Third grade	6.5 (1.8)	5.2 (2.6)
	Sixth grade	4.7 (1.4)	5.8 (1.2)

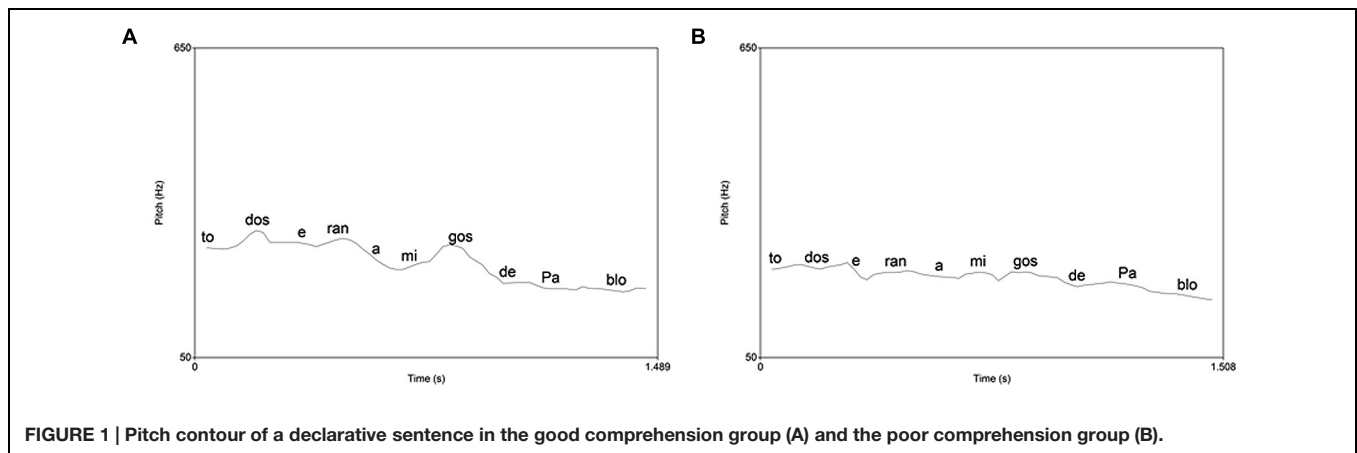


FIGURE 1 | Pitch contour of a declarative sentence in the good comprehension group (A) and the poor comprehension group (B).

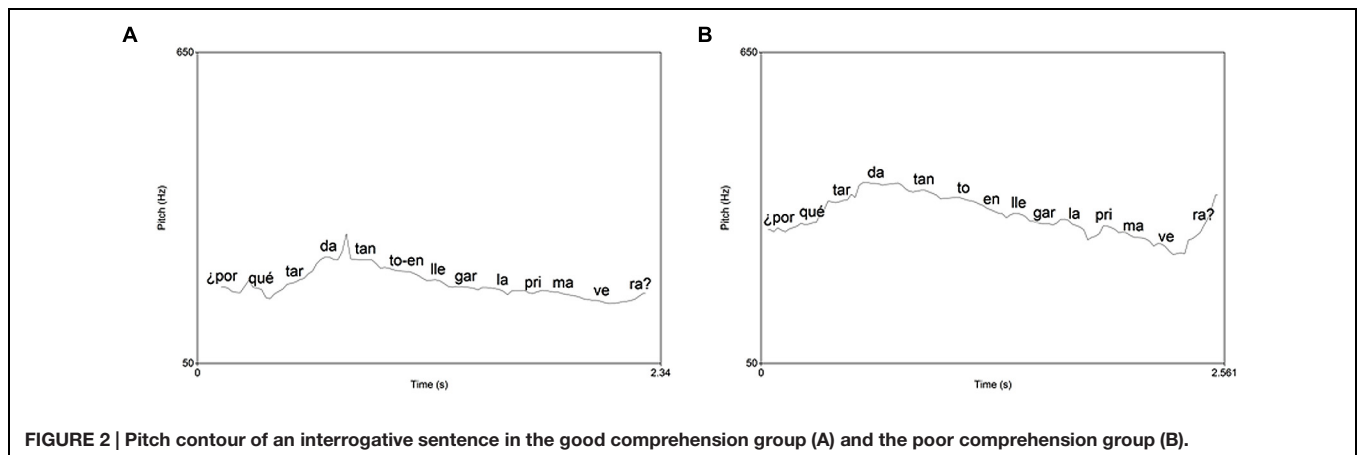


FIGURE 2 | Pitch contour of an interrogative sentence in the good comprehension group (A) and the poor comprehension group (B).

analyzed with SPSS software the results from the Praat analysis comparing the two groups and grades using ANOVAs. We found a significant group effect in the fundamental frequency of syllables in declarative sentences [ $F(1,34) = 4.6, p = 0.038$ ]. In the frequency range of interrogative sentences the interaction between group and grade was also significant [ $F(1,35) = 6.9, p = 0.012$ ]. We made *post hoc* comparisons using the Tukey test, showing that the significant differences were between the two groups of third grade ( $p = 0.048$ ). See Figures 1 and 2 as examples of the effects found in these analyses and Table 4 with the means and SD.

## Words Analyses

Firstly, we analyzed the low frequency words repeated and non-repeated, discarding those misread. SPSS software was used to conduct ANOVAs for comparing groups and grades. No significant effect by groups was found ( $p > 0.05$ ).

Secondly, an ANOVA was performed with the words with different stress and frequency. There was a significant interaction between the number of errors in the words with different lexical frequency by group [ $F(1,36) = 6.4, p = 0.016$ ]. Also an interaction of the mean time for reading high frequency words stressed on the penultimate syllable and group, and grade was found [ $F(1,36) = 5.7, p = 0.022$ ]. *Post hoc*

comparisons using the Tukey HSD test indicated that the mean reading time of children with poor comprehension from third grade was significantly different from the same group of sixth grade ( $p < 0.001$ ). Finally, a similar interaction of the mean time for reading high frequency words stressed on the antepenultimate syllable and group, and grade was found [ $F(1,36) = 4.6, p = 0.038$ ]. The above *post hoc* comparisons also showed that the significant differences were between poor reading comprehension group children in third and sixth grade ( $p = 0.001$ ). See Table 5 for the means and SD.

## DISCUSSION

The aim of this study was to investigate the relationship between comprehension and prosody, both as a part of reading fluency (Rasinski, 2004; Ravid and Mashraki, 2007; Hudson et al., unpublished manuscript). To achieve this objective, we selected two groups of children according to their level of reading comprehension in third and sixth grade. The task consisted of reading aloud a text containing several sentence types and words with different characteristics.

Our results revealed that reading accuracy and reading comprehension are related, as we can see that children with poor

**TABLE 4 | Mean and SD of the principal significant effects in the sentences analyses.**

Significant effect		Poor comprehension <i>M (SD)</i>	Good comprehension <i>M (SD)</i>
F0 of syllables in declarative sentences	Other syllables	221 (7.8)	234 (7.4)
	Last syllable	259 (6.9)	251 (6.5)
F0 range of interrogative sentences	Third grade	196 (61)	157 (34.9)
	Sixth grade	148 (30.8)	194 (63.6)

**TABLE 5 | Mean and SD of the primary significant effects in the word analyses.**

Significant effect		Poor comprehension <i>M (SD)</i>	Good comprehension <i>M (SD)</i>
Number of mistakes	High frequency	0.3 (0.07)	0.25 (0.07)
	Low frequency	1.57 (1.4)	1.02 (1.4)
Mean time for reading HF words stressed on the penultimate syllable	Third grade	0.67 (0.06)	0.62 (0.05)
	Sixth grade	0.54 (0.05)	0.57 (0.04)
Mean time for reading HF words stressed on the antepenultimate syllable	Third grade	0.58 (0.09)	0.51 (0.05)
	Sixth grade	0.46 (0.05)	0.48 (0.05)

reading comprehension made more mistakes in content words than children with good reading comprehension. Also this group was more affected by lexical frequency, since they made a higher number of mistakes in the low frequency words, independently of the stress. A low reading accuracy make children to more misread words and, as Perfetti (1985) stated, readers who fail in word identification will be poorer comprehenders, because of working memory. There are two points of view about the relationship between working memory and reading comprehension, as we could see in the review of this issue made by Van Dyke and Shankweiler (2013). The first one believes that working memory is limited and when it is busy with the decoding not attends to comprehension. The other one related reading comprehension also with high-quality lexical representations. Nevertheless, we have seen that the children with a worse comprehension made more mistakes while reading and also had low scores in the initial subtests of PROLEC-R. Therefore, we could think that one of the causes for poor comprehension could be a low decoding skill that does not allow children to read accurately; as saying by the Simple View of Reading (Hoover and Gough, 1990) decoding is a necessary skill for reading comprehension. It seems clear that children with more reading mistakes show more difficulties to understand when reading because the errors do not allow them process the whole text, but only a part.

On the other hand, better comprehenders had lower reading times in high frequency words with stress on the penultimate and on the antepenultimate syllables than in the same low frequency words. Besides, we found significant differences between grades in the group with poor comprehension; the third-grade children had higher reading times than the sixth-grade children. That did not occur within the group with good reading comprehension, where there was no significant difference between two grades. It could be due to lexical frequency having more weight than the stress place for children with lower reading skills, and as a consequence they read the words with high lexical frequency faster and more accurately. However, this is not what usually

happens, as there is a clear tendency to read the words as stressed on the penultimate syllables and make more mistakes when they are stressed on the last and antepenultimate syllable (Gutiérrez-Palma et al., 1998).

In addition, there is a relationship between pausal intrusions and the understanding of the text, since children with a poor reading comprehension made more inappropriate pauses. This was reported in other studies where children with higher fluency made fewer ungrammatical pauses (Miller and Schwanenflugel, 2006; Benjamin and Schwanenflugel, 2010; Alves et al., 2014); the same relationship appears in adults (Binder et al., 2013), where those with low literacy skills made more sentence intrusions compared to the skilled adult readers. Thus, readers with better decoding and word reading skills paused less frequently than readers who had poorer decoding and word reading skills. In addition, readers who experience fewer word and sentence intrusions had better comprehension abilities. Making many pauses involves an increase of the reading time, which would require greater working memory, as Perfetti (1985) considered that more work memory requirement reduced the number of available resources for understanding. That is, when a reader makes a higher number of pauses more working memory is needed and this means less understanding. This greater number of inappropriate pauses made by those children with poor understanding may be due to a low decoding skill. This group of children seems to have difficulty to decode rightly unfamiliar words (i.e., pseudowords and low frequency words), and this may make them stop inappropriately more often in the middle of the words or before unknown words. But not only were the inappropriate pauses different between groups; also the intersentential pauses before commas were different, since third grade children with poorer comprehension made more pauses before commas than the same group in sixth grade. This was seen by Miller and Schwanenflugel (2006) in a study with third grade children and by Chafe (1988) in his study with adults, where more skilled readers may not feel driven to mark every comma



with a pause. Our results agree with those findings, since the group with poor reading comprehension, who are less skilled, made more pauses before commas than the group with good reading comprehension; besides, within the group with poor comprehension, third-grade children, younger and having lower reading skills, paused more often than children from sixth-grade.

Moreover, a correct prosody involves a proper melodic contour, suitable for every type of sentence. Particularly, in declarative and exclamatory sentences the pitch falls at the end of the sentence, while in the yes-no questions the pitch rises (Miller and Schwanenflugel, 2006). Our results showed that only the children with better reading comprehension made a final declination in declarative sentences. That was found by other authors that related a final declination pitch in declarative sentences in better readers (Ladd, 1984; Wichmann, 1994; Benjamin and Schwanenflugel, 2010). We also found differences in the total range of pitch in the interrogative sentences of third grade, since children with less reading comprehension had a bigger pitch range. We could think that these poor young readers exaggerate the pitch contour when faced with a question mark. It is already known that children are aware of the different linguistic marks, such as exclamatory signs or

quotes (Schwanenflugel et al., 2015), as they modify the tone and intensity when they encounter them. It is not surprising, therefore, that in the early stages of learning to read they tend to exaggerate those prosodic features of certain linguistic marks.

To sum up, the current study provides information about the relationship between prosody and reading comprehension, which is a little-studied field, but of great interest to education, since one of the major problems encountered in the classroom is the low reading comprehension presented by students. Determining the direction of this relationship is still needed. However, we have seen that there are different prosodic features, such as pauses or intonation of declarative and interrogative sentences, which differ according to the levels of understanding of the subject.

## ACKNOWLEDGMENTS

This study was funded by Grant PSI2012-31913 from the Spanish Government and supported by a predoctoral grant from the Foundation for the Promotion of Applied Scientific Research and Technology in Asturias (FICYT).

## REFERENCES

- Adlof, S. M., Catts, H. W., and Little, T. D. (2006). Should the simple view of reading include a fluency component? *Read. Writ.* 19, 933–958. doi: 10.1007/s11145-006-9024-z
- Alves, L. M., Reis, C., and Pinheiro, Â. (2014). Prosody and reading in dyslexic children. *Dyslexia* 21, 35–49. doi: 10.1002/dys.1485
- Benjamin, R. G., and Schwanenflugel, P. J. (2010). Text complexity and oral reading prosody in young readers. *Read. Res. Q.* 45, 388–404. doi: 10.1598/RRQ.45.4.2
- Binder, K. S., Tighe, E., Jiang, Y., Kaftanski, K., Qi, C., and Ardoin, S. P. (2013). Reading expressively and understanding thoroughly: an examination of prosody in adults with low literacy skills. *Read. Writ.* 26, 665–680. doi: 10.1007/s11145-012-9382-7
- Boersma, P., and Weenink, D. (2015). *Praat: Doing Phonetics by Computer [Computer program]* 5.4.12. Available at: <http://www.praat.org/>
- Calet, N., Defior, S., and Gutiérrez-Palma, N. (2015). A cross-sectional study of fluency and reading comprehension in Spanish primary school children. *J. Res. Read.* 38, 272–285. doi: 10.1111/1467-9817.12019
- Chafe, W. (1988). "Linking intonation units in spoken english," in *Clause Combing Discourse and Grammar*, eds J. Haiman and S. A. Thompson (Amsterdam: John Benjamins).
- Clay, M. M., and Imlach, R. H. (1971). Juncture, pitch, and stress as reading behavior variables. *J. Verb. Learn. Verb. Behav.* 10, 133–139. doi: 10.1016/S0022-5371(71)80004-X
- Cuetos, F., and Barbón, A. (2006). Word naming in Spanish. *Euro. J. Cogn. Psychol.* 18, 415–436. doi: 10.1080/13594320500165896
- Cuetos, F., Rodríguez, B., Ruano, E., and Arribas, D. (2007). *Prolec-R. Bateria de Evaluación de los Procesos Lectores Revisada*. Madrid: TEA ediciones.
- Davies, R., Wilson, M., Cuetos, F., and Burani, C. (2014). Reading in Spanish and Italian: effects of age of acquisition in transparent orthographies? *Q. J. Exp. Psychol. (Hove)* 67, 1808–1825. doi: 10.1080/17470218.2013.872155
- Dowhower, S. L. (1991). Speaking of prosody: fluency's unattended bedfellow. *Theor. Into Pract.* 30, 165–175. doi: 10.1080/00405849109543497
- Fischer-Baum, S., Dickson, D. S., and Federmeier, K. D. (2014). Frequency and regularity effects in reading are task dependent: evidence from ERPs. *Lang. Cogn. Neurosci.* 29, 1342–1355. doi: 10.1080/23273798.2014.927067
- González-Trujillo, M. C., Calet, N., Defior, S., and Gutiérrez-Palma, N. (2014). Scale of reading fluency in Spanish: measuring the components of fluency/Escala de fluidez lectora en español: midiendo los componentes de la fluidez. *Estud. Psicol.* 35, 104–136. doi: 10.1080/02109395.2014.893651
- Gutiérrez-Palma, N., Palma, A., and Santiago, J. (1998). On the role of lexical stress in visual word recognition. *Poster Presented at the XI Congress of the European Society for Cognitive Psychology*, Jerusalem. 13–17.
- Hoover, W. A., and Gough, P. B. (1990). The simple view of reading. *Read. Writ.* 2, 127–160.
- Kameenui, E. J., Fuchs, L., Francis, D. J., Good, R. III, O'Connor, R. E., Simmons, D. C., et al. (2006). The adequacy of tools for assessing reading competence: a framework and review. *Educ. Res.* 35, 3–11. doi: 10.3102/0013189X035004003
- Kaminski, R. A., and Good, R. H. (1998). Toward a technology for assessing basic early literacy skills. *Sch. Psychol. Rev.* 25, 215–227.
- Kim, Y. S., and Wagner, R. K. (2015). Text (Oral) Reading fluency as a construct in reading development: an investigation of its mediating role for children from grades 1 to 4. *Sci. Stud. Read.* 19, 224–242. doi: 10.1080/10888438.2015.1007375
- Klauda, S. L., and Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *J. Educ. Psychol.* 100, 310–321. doi: 10.1037/0022-0663.100.2.310
- Kuhn, M. R., and Stahl, S. (2003). Fluency: a review of developmental and remedial practices. *J. Educ. Psychol.* 95, 3–21. doi: 10.1037/0022-0663.95.1.3
- Ladd, B. (1984). Declination: a review and some hypotheses. *Phonol. Yearbook* 1, 53–74. doi: 10.1017/S0952675700000294
- 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(Pt 2), 269–287. doi: 10.1348/000712602162580
- Martínez, J. A., and García, M. E. (2004). *Diccionario de Frecuencias del Castellano Escrito en Niños de 6 a 12 Años*. Salamanca: Universidad Pontificia de Salamanca, Servicio de Publicaciones.
- Miller, J., and Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *J. Educ. Psychol.* 98, 839–843. doi: 10.1037/0022-0663.98.4.839
- Miller, J., and Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Read. Res. Q.* 43, 336–354. doi: 10.1598/rrq.43.4.2
- Monaghan, P., and Ellis, A. W. (2010). Modeling reading development: cumulative, incremental learning in a computational model of word naming. *J. Mem. Lang.* 63, 506–525. doi: 10.1016/j.jml.2010.08.003
- Muncer, S. J., Knight, D., and Adams, J. W. (2014). Bigram frequency, number of syllables and morphemes and their effects on lexical decision and word naming. *J. Psycholinguist. Res.* 43, 241–54. doi: 10.1007/s10936-013-9252-8

- National Institute of Child Health, and Human Development (2000). *Report of the National Reading Panel. Teaching Children to Read: An Evidence-based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction: Reports of the Subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Perfetti, C. A. (1985). *Reading Ability*. New York, NY: Oxford University Press.
- Petscher, Y., and Kim, Y. S. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *J. Sch. Psychol.* 49, 107–29. doi: 10.1016/j.jsp.2010.09.004
- Rasinski, T. V. (2004). *Assessing Reading Fluency*. Honolulu, HI: Pacific Resources for Education and Learning.
- Rasinski, T., Rikli, A., and Johnston, S. (2009). Reading fluency: more than automaticity? More than a concern for the primary Grades? *Lit. Res. Instr.* 48, 350–361. doi: 10.1080/19388070802468715
- Ravid, D., and Mashraki, Y. E. (2007). Prosodic reading, reading comprehension and morphological skills in Hebrew-speaking fourth graders. *J. Res. Read.* 30, 140–156. doi: 10.1111/j.1467-9817.2007.00340.x
- Richards, M. (2000). Be a good detective: solve the case of oral reading fluency. *Read. Teacher* 53, 534–539.
- Rouibah, A., Taft, M., and Albert, F. (2000). “French sublexical unit priming in pronunciation and lexical decision tasks,” in *Proceeding of the 41th Annual Meeting of the Psychonomic Society, Abstracts of the Psychonomic Society*, New-Orleans, LA.
- Schwanenflugel, P. J., Hamilton, A. M., Wisenbaker, J. M., Kuhn, M. R., and Stahl, S. A. (2004). Becoming a fluent reader: reading skill and prosodic features in the oral reading of young readers. *J. Educ. Psychol.* 96, 119–129. doi: 10.1037/0022-0663.96.1.119
- Schwanenflugel, P. J., Westmoreland, M. R., and Benjamin, R. G. (2015). Reading fluency skill and the prosodic marking of linguistics focus. *Read. Writing Interdiscip. J.* 28, 9–30. doi: 10.1007/s11145-013-9456-1
- Seymour, P. H., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94(Pt 2), 143–174. doi: 10.1348/000712603321661859
- Taft, M. (1979). Lexical access-via an orthographic code: the basic orthographic syllabic structure (BOSS). *J. Verb. Learn. Verb. Behav.* 18, 21–39. doi: 10.1016/S0022-5371(79)90544-9
- Van Dyke, J. A., and Shankweiler, D. (2013). “From verbal efficiency theory to lexical quality: the role of memory processes in reading comprehension,” in *Reading: From Words to Multiple Texts*, eds A. Britt, S. Goldman, and J.-F. Rouet (London: Routledge), 115–131.
- Wichmann, A. (1994). F0 troughs and prosodic phrasing: an investigation into the linguistic information contained in a speaker’s base line when reading. *Hermes* 13, 205–212.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Álvarez-Cañizo, Suárez-Coalla and Cueto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Preliminary validation of FastaReada as a measure of reading fluency

Zena Elhassan<sup>1</sup>, Sheila G. Crewther<sup>1\*</sup>, Edith L. Bavin<sup>1</sup> and David P. Crewther<sup>2</sup>

<sup>1</sup> Department of Psychology and Counselling, La Trobe University, Bundoora, VIC, Australia, <sup>2</sup> Brain Sciences Institute, Swinburne University, Hawthorn, VIC, Australia

## OPEN ACCESS

### Edited by:

Giseli Donadon Germano,  
Universidade Estadual Paulista, Brazil

### Reviewed by:

Adriana De Souza Batista Kida,  
Universidade Estadual Paulista "Julio  
de Mesquita Filho", Brazil  
Adriana Marques De Oliveira,  
Universidade Estadual Paulista "Julio  
de Mesquita Filho", Brazil

### \*Correspondence:

Sheila G. Crewther  
s.crewther@latrobe.edu.au

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 03 June 2015

**Accepted:** 09 October 2015

**Published:** 27 October 2015

### Citation:

Elhassan Z, Crewther SG, Bavin EL  
and Crewther DP (2015) Preliminary  
validation of FastaReada  
as a measure of reading fluency.  
Front. Psychol. 6:1634.  
doi: 10.3389/fpsyg.2015.01634

Fluent reading is characterized by speed and accuracy in the decoding and comprehension of connected text. Although a variety of measures are available for the assessment of reading skills most tests do not evaluate rate of text recognition as reflected in fluent reading. Here we evaluate FastaReada, a customized computer-generated task that was developed to address some of the limitations of currently available measures of reading skills. FastaReada provides a rapid assessment of reading fluency quantified as words read per minute for connected, meaningful text. To test the criterion validity of FastaReada, 124 mainstream school children with typical sensory, mental and motor development were assessed. Performance on FastaReada was correlated with the established Neale Analysis of Reading Ability (NARA) measures of text reading accuracy, rate and comprehension, and common single word measures of pseudoword (non-word) reading, phonetic decoding, phonological awareness (PA) and mode of word decoding (i.e., visual or eidetic versus auditory or phonetic). The results demonstrated strong positive correlations between FastaReada performance and NARA reading rate ( $r = 0.75$ ), accuracy ( $r = 0.83$ ) and comprehension ( $r = 0.63$ ) scores providing evidence for criterion-related validity. Additional evidence for criterion validity was demonstrated through strong positive correlations between FastaReada and both single word eidetic ( $r = 0.81$ ) and phonetic decoding skills ( $r = 0.68$ ). The results also demonstrated FastaReada to be a stronger predictor of eidetic decoding than the NARA rate measure, with FastaReada predicting 14.4% of the variance compared to 2.6% predicted by NARA rate. FastaReada was therefore deemed to be a valid tool for educators, clinicians, and research related assessment of reading accuracy and rate. As expected, analysis with hierarchical regressions also highlighted the closer relationship of fluent reading to rapid visual word recognition than to phonological-based skills. Eidetic decoding was the strongest predictor of FastaReada performance (16.8%) followed by phonetic decoding skill (1.7%). PA did not make a unique contribution after eidetic decoding and phonetic decoding skills were accounted for.

**Keywords:** FastaReada, reading fluency, reading development, assessment, automaticity, visual word recognition, phonological awareness

## INTRODUCTION

Fluent reading is characterized by speed and accuracy in the decoding of connected text (Fuchs et al., 2001). Although reading fluency is regarded as a key component in the maturation of reading skill (e.g., Samuels, 2006), formal consideration of the construct remains limited (e.g., Kame'enui and Simmons, 2001; Kuhn et al., 2010; Valencia et al., 2010). Furthermore, few

studies have attempted to clarify the factors facilitating rate and mode (i.e., visual or eidetic versus auditory or phonetic) of word decoding in fluent reading of comprehensible text. In the current study we assessed FastaReada (Hecht et al., 2004), a computer-based measure of reading fluency by comparing children's performance on the measure with scores for reading accuracy, comprehension, and rate on a well-established test of text reading ability, the Neale Analysis of Reading Ability (NARA; Neale, 1999) and single word reading on the dyslexia Determination Test (DDT; Griffin and Walton, 1987). To demonstrate criterion-related validity, evidence that FastaReada performance assesses the core features of fluency that is strong, positive correlations with established measures of text reading rate, accuracy and comprehension was sought (Petscher and Kim, 2011). Before discussing the new measure, we discuss our understanding of fluent reading and factors that influence it.

Fluent reading is a multifaceted cognitive process that is usually considered to be dependent on the development of numerous endogenous skills such as phonological awareness (PA; e.g., Ziegler and Goswami, 2005), letter knowledge (e.g., Blaklock, 2004), visual recognition (e.g., Sereno and Rayner, 2003), attention (e.g., Kinsey et al., 2004), working memory (e.g., Daneman and Carpenter, 1980), naming speed (e.g., Logan, 1997) and speed of processing (e.g., Breznitz and Misra, 2003). Exogenous factors, such as text characteristics, purpose for reading and reading topic have also been shown to contribute to reading fluency (refer to Hosp and Suchey, 2014 for a review).

Following the early work of Vellutino (1977, 1979; e.g., learning to read has most often been associated with and attributed to competence in PA). PA is usually defined as the ability to deconstruct spoken words into distinctive sounds, or phonemes (the distinctive sounds of a language), syllables, and onsets and rimes (e.g., for 'bat': /b/ onset and [æt] rime) (Liberman, 1971; Treiman and Zukowski, 2013). When alphabetic orthographic systems have a high degree of consistency between grapheme, or letter and phoneme correspondence (Ehri, 1992), phoneme-grapheme knowledge, which is underpinned by PA, provides learner readers with a basic strategy for decoding printed text into its spoken form (Castles and Nation, 2006). However, dependence on this slow and laborious approach to reading, which is termed phonetic decoding, gradually decreases as learner readers acquire orthographic and vocabulary skills that facilitate rapid visual recognition of printed words and subsequent fluency (Ehri and Wilce, 1980, 1985; Thomson et al., 2006; Vellutino et al., 2007). Furthermore, alphabetic systems differ in their orthographic complexity; for example, many English words are irregular words that do not conform to typical grapheme-phoneme mapping rules (e.g., *yacht* and *colonel*). Such words necessitate visual recognition for accurate decoding (Boder, 1973; Castles and Coltheart, 1993). Whilst the relationship between PA and reading skills for list words has been widely studied (e.g., Wagner and Torgesen, 1987; Del Campo et al., 2015), the role of PA in fluent reading is not well researched.

Word recognition occurs when the visual representation of a word corresponds with a stored phonological representation in the mental lexicon (Taft, 1986). With practice, whole word recognition, and hence single word and text reading becomes progressively faster, seemingly automatic and reflex-like (Meyer and Felton, 1999; Hecht et al., 2004; Laycock and Crewther, 2008). Fluency in reading is less demanding of cognitive resources than conscious decoding and therefore frees up attentional stores for higher level processing, that is, comprehension (La Berge and Samuels, 1974; Perfetti et al., 1988). In support, numerous studies have shown strong, positive correlations between reading rate and comprehension (Breznitz, 1987; Jenkins et al., 2003; O'Connor et al., 2010). Thus any new tool for assessing reading must be shown to correlate well with assessments that test accuracy of comprehension.

Most available tools for assessments of reading ability operationalize reading fluency as words correctly identified per minute (WCPM) minus errors such as mispronunciations, substitutions, omissions and insertions (Valencia et al., 2010). The Gray Oral Reading Test (GORT) and the Kaufman Test of Educational Achievement (KTEA) provide specific scores of "reading fluency" via a WCPM score. The KTEA utilizes single word list stimuli rather than connected text, which reduces its criterion validity, as fluent reading is often used to acquire meaningful information from connected text. A further disadvantage of single-word reading rate measures is that they are not effective for the identification of children with specific reading comprehension deficits whereas rate of reading for connected text has been shown to differentiate children with and without comprehension deficits (Cutting et al., 2009). While both the GORT and KTEA are marketed as providing measures of fluency, only the GORT and the NARA utilize connected text stimuli. The NARA does not claim to measure fluency; however, its rate score is representative of words read accurately per minute *out loud*, making it akin to the GORT as a measure of fluency. The GORT, KTEA, and NARA all provide a measure of reading comprehension.

Further disadvantages of the GORT, KTEA, and NARA are related to their availability to educators. Such assessments are restricted to professionals trained in the administration and interpretation of norm-referenced standardized tests. Furthermore, the standardized nature of the reading material used in such assessments is likely to increase repeated administration effects (i.e., practice effects) if used frequently. These factors render such tests as inappropriate for regular monitoring of reading skill development, an important strategy for the development of personalized learning plans (Deno, 2003). Long term established benefits of regular monitoring of reading progress include improved learning outcomes, enhanced educator decision-making and increased student awareness of their own performance (e.g., Fuchs and Deno, 1991; Fuchs and Fuchs, 2002).

In order to enable educators to confidently monitor student reading development progress we have developed a customized



computer-generated task called FastaReada. FastaReada has been designed to provide a quick and reliable measure of rate and accuracy of text reading. The defining feature of this task that sets it apart from other WCPM measures is that it utilizes a maximum-likelihood parameter estimation by sequential testing (PEST) testing method to establish the threshold exposure time required to decode short pieces of text (six words). Testing begins with a long exposure time that typically developing students can easily verbalize. With each correct response verbalized, the exposure time becomes shorter, encouraging fluent readers to read silently whilst text is exposed, and then to repeat the words after they disappear. This technique requires encoding of words visualized prior to verbalization, ensuring that FastaReada can also test aspects of working memory, and the cognitive speed of reading by reducing the impact of motor limits on verbal reaction times (Swanson et al., 2009).

The current research compared FastaReada performances with reading accuracy, rate, and comprehension scores on the well-established test of reading ability, the NARA to meet the requirement of criterion-related validity that is, strong, positive correlations between all variables. To address shortcomings in the literature, the current research also aimed to examine the contribution of PA, as measured by the Comprehensive Test of Phonological Processing, (CTOPP) to fluent reading. In addition, the contributions to fluent reading from phonetic decoding skills, as measured by the Pseudoword Decoding subtest of the Wechsler Individual Achievement Test, and relative ability to eidetically recognize words compared need phonetic decoding, as measured by the Dyslexia Determination Test, were also investigated. It was hypothesized that:

- (1) FastaReada scores would be strongly associated with NARA measures of reading accuracy, comprehension and rate.
- (2) Visual word recognition would be a stronger predictor of FastaReada performance than phonetic decoding skill and PA.
- (3) The association between FastaReada scores and visual word recognition would be stronger than the association between NARA rate scores and visual word recognition.

## MATERIALS AND METHODS

### Participants

This study was approved by the La Trobe University Human Ethics Committee (FSTE HEC 13/R22). Consent to collect data from schools was also provided by the Victorian State Department of Education (2012\_001425) and Catholic Education Melbourne (GE12/0009 1765). One hundred and twenty-nine children between the ages of 9–12 years were recruited from 3 year levels (Year 4–6) through four mainstream schools in the North East Metropolitan region of Melbourne, Australia. The schools covered regions of high and low socioeconomic conditions. Participant information and consent forms were dissemination to parents and legal guardians of children in the

target year levels. Every child who returned a signed consent form was permitted to undergo the entire battery of tests to avoid leaving children with a sense of exclusion. However, for inclusion in data analysis participants required a score above the 10th percentile on the Raven's Coloured Progressive Matrices (RCPM: a test of non-verbal reasoning ability; Raven et al., 1998), adequate or adequately corrected vision and hearing, and typical sensory, mental and motor development. Three children were excluded from analysis for scoring at or below the 10th percentile on the RCPM. A further two were excluded on the basis of teacher report of confirmed or suspected neurodevelopmental disorder. The final number of participants was 124 (see **Table 1** for demographics).

### Materials

#### FastaReada (Hecht et al., 2004)

FastaReada is a customized computer-generated task designed using VPixx (www.VPixx.com) that measures reading fluency. An excerpt from a contemporary novel, which appeals to children between the ages of 9–12 years of age (permission received from Penguin Group) is presented in narrative order, six words at a time. The presentation time for the group of words presented (Lucida Grande font, 60 pt) was controlled via the PEST adaptive staircase algorithm based on a maximum-likelihood threshold estimation. Children were asked to read the stimulus out loud as accurately as possible. The investigator indicated accurate or inaccurate decoding at the end of each trial. Prior to assessment with FastaReada children were warned that the duration of stimuli presentation would eventually become so short that they would not be able to read all six words out loud. They were encouraged to attempt each trial in spite of the brief exposure time.

#### Neale Analysis of Reading Ability–Third Edition (NARA-3; Neale, 1999)

The NARA-3 is commonly used in school and clinical settings as a standardized measure of reading achievement and diagnostic test. It provides objective measures of reading accuracy, reading comprehension, and reading rate in children aged from 6 to 12 years and over. Administration takes approximately 20 min (Neale, 1999). The NARA-3 was administered according to the standard procedure for testing, as outlined in the NARA-3 manual (Neale, 1999).

In summary, children were instructed to read a series of prose passages presented in book form and answer questions about each passage at its conclusion. Each passage was accompanied

**TABLE 1 | Mean, minimum and maximum age of children (in years and months), and number of each sex in each year level.**

	Mean age	Minimum age	Maximum age	Male	Female
Year 4	9;11	9;1	10;5	17	25
Year 5	10;10	10;2	11;8	18	23
Year 6	12;0	11;4	12;10	20	21
<b>All year levels</b>	10;11	9;1	12;0	55	69

by a line drawing that was intended to set the scene rather than to provide detail. The investigator corrected and recorded the number of errors, including mispronunciations, substitutions (i.e., real words used instead of the word in the passage), refusals (i.e., child pauses for approximately 4–6 s and does not make an attempt at the word), additions (i.e., child inserts words or part of words into the passage), omissions (i.e., child omits words from the passage), and reversals (e.g., child says ‘no’ for ‘on’). The investigator also recorded the time the child took to read each passage and marked the child’s answers to questions as correct or incorrect. There were six passages in total, which were presented in order of increasing difficulty. Testing was discontinued after the child reached the ceiling for reading errors in a passage (16 errors for passages 1–5 or 20 errors for passage 6). Separate scores for accuracy, comprehension and reading rate were obtained.

The NARA-3 has been shown to be a reliable and valid tool for the assessment of accuracy, rate and comprehension of oral reading skills. Reliability results for the NARA-3 ranged from moderate to high levels of internal consistency across groups based on years of schooling (0.91–0.96 for accuracy, 0.71–0.96 for comprehension, and 0.73–0.96 for rate (Neale, 1999). The assessment has been shown to have high content and face validity for the construct of oral reading (Neale, 1999). Additionally, it has been shown to have criterion-related validity through its significant correlations with other tests of reading skills (e.g., Moorehouse and Yule, 1974) and through its efficacy at predicting future reading ability (McKay, unpublished as cited in Neale, 1999). Finally, the positive correlation between score and years of schooling provides evidence of construct related validity (Neale, 1999).

### **Dyslexia Determination Test (DDT; Griffin and Walton, 1987)**

The DDT is a diagnostic assessment tool that is used to identify the nature, type, and severity of an individual’s learning difficulties (Wesson, 1993). The DDT is designed for students between Year 2 to Year 12 levels and consists of three sections that assess single word reading, writing, and spelling abilities of children (Wagner et al., 1999). The current study utilized the DDT decoding subtest to identify the degree of visual recognition and phonetic decoding strategies used by each child when reading word lists. The DDT was administered according to the standard procedure for testing (Wagner et al., 1999).

Children were asked to commence orally decoding the list words from the initial list (i.e., the pre-primer words) rather than the suggested two to three levels below their year level in order to avoid frustration and to assist with building confidence. The items on the list alternate between phonetically irregular words (i.e., requiring visual recognition for accurate decoding) and phonetically regular words (i.e., conforming to English letter-sound rules). In line with the standard DDT procedure, words read correctly within 2 s were marked as eidetic (i.e., visually recognized) on the DDT record form as the rapid response is indicative of visual recognition. Words read correctly after a delay of more than 2 s but within 10 s were considered to require phonetic

decoding indicating the use of phonics, syllabication and/or structural analysis in word decoding. Words that were not read within 10 s, read incorrectly or not attempted were marked as unknown.

### **Pseudoword Decoding: Wechsler Individual Achievement Test–Second Edition (WIAT-II)**

The Pseudoword Decoding subtest of the WIAT-II was used to measure phonetic decoding skills. It consists of 54 non-word items, all of which conform to letter-sound rules of regular English words, making the task similar to encountering and decoding unfamiliar words. The investigator administered the task in line with the general assessment procedure (Wechsler, 2007).

In summary, the investigator asked the children to read each item on the Pseudoword Card, from left to right. All children began at the same starting point and the discontinue rule was met once seven consecutive incorrect responses were made.

The WIAT-II is a well-established test of individual achievement. The most reliable and valid measures attained by the WIAT-II are the composite scores; however, the degrees of reliability and validity across individual tests have been shown to be adequate. The Pseudoword Decoding subtest has been shown to be a reliable measure of non-word decoding skill, with high-level inter-item reliability (0.89–0.98) and test–retest stability (0.93) across ages 6–19 years. Evidence of construct- and criterion-related validity has also been demonstrated across the subtests (Wechsler, 2007).

### **Phonological Awareness: Comprehensive Test of Phonological Processing**

Participants completed the two PA subtests available for their age group on the CTOPP. The results from these two tasks formed the PA composite score. The subtests assessed elision (the exclusion of one or more sounds from a word) and sound blending (the ability to build whole words by blending individual sounds together). Both PA subtests were administered according to the standard procedure (Wagner et al., 1999).

The PA subtests required the investigator to provide feedback for practice items and the first three test items. Each item could be repeated one additional time if requested by the child. Testing was discontinued following three consecutive incorrect responses. For the elision subtest the investigator asked the child to say a compound-word. After the word was verbalized, the investigator asked the child to say the word again without one of the segments (e.g., “Say *steamboat* without saying *boat*”). For the sound blending subtest children were instructed to listen carefully as words were voiced in small parts, one part at a time, and then to put the parts together to verbalize the whole word (e.g., “*What word do these sounds make when you put them together c-o-m-p-u-t-e-r?*”). Data analysis was conducted with raw composite scores.

The CTOPP has been shown to be a reliable and valid tool for the assessment of PA. Wagner et al. (1999) demonstrated moderate to high internal consistency for all subtests across groups based on age (specifically 0.81 to 0.92 for Elision and

0.78 to 0.89 for Sound Blending). Reliability was also high for time sampling and inter-scorer differences (Wagner et al., 1999). Strong correlations have been demonstrated between the PA composite of the CTOPP and the Lindamood Auditory Conceptualization Test (Wagner et al., 1999), Woodcock Reading Mastery Tests – Revised (Wagner et al., 1994, 1997), as well as the Test of Word Reading Efficiency (Wagner et al., 1999), providing support for criterion-prediction validity. Additionally, construct validity of the CTOPP is demonstrated by the positive correlation between age and score, and the test items are sufficiently correlated for the verification of content validity (Wagner et al., 1999).

### Raven's Colored Progressive Matrices (RCPM) Test

The RCPM (Raven et al., 1998) was used to provide a standardized, untimed, non-verbal measure of general intelligence through the assessment of non-verbal reasoning ability. The RCPM has been norm-referenced in numerous countries, including Australia, and earlier work has shown that the RCPM is an appropriate measure for typically developing children aged 5 to 11 years, as well as children with reading and/or learning disorders (Cotton et al., 2005b).

The RCPM consists of three sets of 12 colored multiple-choice items that gradually increase in complexity (A, Ab, B). Each item is presented on an A4 sized sheet of paper and consists of an incomplete matrix. Children were asked to identify or point to one of six figures positioned below the rectangle that would correctly complete the pattern. A score of one point was rewarded for each correct answer, whilst incorrect answers scored a zero. The scores were tallied upon completion of the task to provide an overall raw score. Raw scores were converted to percentile scores to rank non-verbal intelligence on the basis of chronological age.

The RCPM has been demonstrated to have good test-retest reliability at  $r = 0.80$  (Raven et al., 1998) and high internal consistency ( $r = 0.89$ ), with minimal variation across age levels (Cotton et al., 2005b).

### Procedure

Testing was conducted over three sessions that ran for approximately 30 min each, in order to reduce disruptions to classroom learning. The testing sessions were run during school hours, in a quiet room within the child's school. The order of tests was determined in an order that would promote interest and reduce fatigue (i.e., cognitively demanding and paper-based tests were limited in each session and computer-based tests were administered toward the end of each session to act as an incentive).

Prior to the commencement of each session, each child was asked "Would you like to play some paper and computer games with us?" and encouraged to request breaks or tell the investigator if they wanted to stop participating. All children recruited stated that they wanted to participate and there were no requests made for breaks or termination of participation. Children were praised for their performances at the conclusion of each session and were encouraged to choose a "thank you gift" from a box of novelty stationary items.

### Statistical Procedures

Data was screened for accuracy of entry, missing values and violations of the assumptions of statistical tests, prior to statistical analysis using Statistical Package for Social Scientists (IBM SPSS Statistics 22). The data set was deemed to be accurate and free from missing values. Preliminary analyses of all data were conducted to assess the assumptions of homoscedasticity, linearity and homogeneity of variance. The frequency distribution of each variable was assessed for violations of normality using standardized indices ( $z$ ) of skewness and kurtosis with a conservative criterion of  $\alpha = 0.001$ ; half the variables were considered close to normal, with skewness and kurtosis values falling between  $-6.56$  and  $+9.47$ . Outliers identified for the variable *FastaReada* were rescored to the next lowest score identified to reduce influence on remaining data (Tabachnick and Fidell, 2013). A square root transformation was then applied to the *FastaReada* variable. Square root transformations resulted in substantial improvement for variables in violation of normality, including *FastaReada*, PA, eidetic decoding, Pseudoword decoding, and NARA-3 accuracy. Reflected transformations were applied to NARA-3 accuracy, eidetic decoding, pseudoword decoding, and PA variables (Tabachnick and Fidell, 2013). No interactions were found between the variables.

Pearson product moment correlations and hierarchical regressions were used to explore the data. Correlation coefficients ( $r$ ) are reported to quantify the degree and direction of the relationships between variables with 0.10–0.29 considered a weak relationship, 0.30–0.49 a medium relationship, and 0.50–1.0 a strong relationship (Cohen, 1988, pp. 79–81). Hierarchical regression was used to explore the proportion of variance in the dependent variables that could be accounted for by one or more independent variables (i.e., how well the independent variables predicted the dependent variable). Change in multiple correlation coefficient squared ( $R^2$ ) values were reported on a range from 0 to 100% to indicate the proportion of variance that was accounted for by each set of independent variables. Squared semi-partial correlations ( $Sr^2$ ) were used to quantify the unique contributions of individual independent variables.

## RESULTS

### Hypothesis One: Relationship between FastaReada Scores and NARA-3

Bivariate correlations were conducted on *FastaReada* scores and the scores obtained on the three NARA-3 variables. The relationships between *FastaReada* performance, and performances on the accuracy, rate, and comprehension subtests were investigated using Pearson product-moment correlation coefficient. Results of the analysis are shown in **Table 2**. Higher NARA-3 accuracy and rate scores were strongly associated with higher *FastaReada* scores across all year levels tested ( $r = 0.79$ – $0.85$  for accuracy and  $r = 0.61$ – $0.82$  for rate). Higher NARA-3 comprehension scores were strongly associated with higher *FastaReada* scores for children in Years 5 ( $r = 0.70$ ) and

**TABLE 2 | Correlations between FastaReada Scores with NARA-3 accuracy, comprehension, and rate subtests for each year level.**

	NARA-3		
	Accuracy	Comprehension	Rate
Grade 4	0.79**	0.47**	0.61**
Grade 5	0.81**	0.70**	0.82**
Grade 6	0.85**	0.64**	0.78**
<b>All Year Levels</b>	0.83**	0.63**	0.75**

NARA-3, Neale Analysis of Reading Ability – Third Edition. \*\* $p < 0.01$ .  $N = 124$ , Grade 4  $n = 42$ , Grade 5  $n = 41$ , Grade 6  $n = 41$ .

6 ( $r = 0.64$ ). Higher Year 4 NARA-3 comprehension scores were moderately positively correlated with rapid and accurate performance on FastaReada ( $r = 0.47$ ).

## Hypothesis Two: Relationships between FastaReada and Eidetic and Phonological Decoding

A preliminary correlation matrix was run in order to assess the associations between visual word recognition, FastaReada performance, phonetic decoding, and PA. The relationships between FastaReada performance and performance on the remainder of the variables were investigated using Pearson product-moment correlation coefficient. Results of the analysis are shown in **Table 3**. The results revealed strong positive correlations between FastaReada scores and scores on the visual word recognition component of the DDT across year levels ( $r = 0.77$ – $0.82$ ). Phonetic decoding ability and performance on FastaReada were also strongly positively correlated ( $r = 0.59$ – $0.76$ ). The correlation between FastaReada performance and PA was moderately positively correlated in Year 5 ( $r = 0.48$ ) and 6 children ( $r = 0.44$ ). However, PA was not shown to be associated with FastaReada performance at a statistically significant level in Year 4 children ( $r = 0.20$ ). A weak to moderate negative correlation was documented between the level of phonetic decoding on the DDT and FastaReada performance ( $r = -0.32$  to  $-0.19$ ). When all year levels were combined, a moderate positive association was shown between FastaReada performance and PA ( $r = 0.37$ ).

**TABLE 3 | Correlations between FastaReada scores and scores on tests of decoding mode, phonetic decoding skill and phonological awareness for each year level.**

	DDT		WIAT-II	CTOPP
	Eidetic	Phonetic	Phonological decoding	Phonological awareness
Grade 4	0.82**	-0.20**	0.59**	0.20**
Grade 5	0.82**	-0.32**	0.76**	0.48**
Grade 6	0.77**	-0.19**	0.71**	0.44**
<b>All year Levels</b>	0.81**	-0.25**	0.68**	0.37**

WIAT-II, Wechsler Individual Achievement Test – Second Edition; CTOPP, comprehensive test of phonological processing; DDT, Dyslexia Determination Test. \*\* $p < 0.01$ .  $N = 124$ , Grade 4  $n = 42$ , Grade 5  $n = 41$ , Grade 6  $n = 41$ .

Hierarchical regression analysis was conducted to ascertain the contributions of PA, phonetic decoding, and eidetic decoding skills to reading fluency. The regression controlled for the contribution of age and non-verbal reasoning ability to the variance in FastaReada scores in the first step. The second step explored the variance in FastaReada scores attributable to PA and phonetic decoding abilities. The eidetic decoding variable was entered in the final step in order to control for the contributions of phonological-based skills. The results are presented in **Table 4**.

The results revealed that age and non-verbal reasoning (Step 1) accounted for 18% of the variance in FastaReada performance,  $F(2,121) = 13.32$ ,  $p < 0.001$ . The addition of the phonological-based skills in Step 2 contributed significantly to the regression model, explaining a further 33.4% of the variance in FastaReada performance,  $F(3,119) = 40.99$ ,  $p < 0.001$ . Finally, the introduction of eidetic decoding skills explained a further 16.6% of the variance in FastaReada performance. The visual word recognition strategy of decoding was the strongest unique contributor to FastaReada performance (16.8%). Phonetic decoding skill also contributed unique variance of 1.7%. Age, non-verbal reasoning and PA did not make unique contributions in the final model. Together, the five independent variables accounted for 68.1% of the variance in FastaReada performance.

## Hypothesis Three: Relationship between FastaReada and Visual Word Recognition and NARA-3 Reading Rate

Bivariate correlations were conducted on visual word recognition scores and the scores obtained on FastaReada and the NARA-3 rate subtest. The relationships between visual word recognition strategy utilization and scores on FastaReada and the NARA-3 rate subtest were investigated

**TABLE 4 | Hierarchical regression results of non-verbal reasoning, age in years, visual word recognition, phonetic decoding skill, and phonological awareness predicting overall FastaReada performance.**

Predictor	Final summary			Step summary	
	Beta	r	Sr	$\Delta R^2$	p
Step 1				0.180	<0.001
Age (years)	0.26	0.31	0.26		
Non-verbal reasoning	0.30	0.34	0.30		
Step 2				0.334	<0.001
Age (years)	0.16	0.31	0.15		
Non-verbal reasoning	0.15	0.34	0.15		
Phonological awareness	0.03	0.37	0.02		
Phonetic decoding skills	0.59	0.68	0.48		
Step 3				0.166	<0.001
Age (years)	0.03	0.31	0.03		
Non-verbal reasoning	0.07	0.34	0.07		
Phonological awareness	0.00	0.37	0.00		
Phonetic decoding	0.20	0.68	0.13		
Eidetic decoding	0.63	0.81	0.41		

$N = 124$ .



using Pearson product-moment correlation coefficient. The results revealed strong positive correlations between both FastaReada and NARA-3 rate scores and visual word recognition strategy utilization. Visual word recognition was more closely associated with FastaReada performance,  $r(122) = 0.81$ ,  $p < 0.001$ , than with NARA-3 rate performance,  $r(122) = 0.73$ ,  $p < 0.001$ .

Further analysis using hierarchical regression controlled for the contribution of age and non-verbal reasoning ability to the variance in eidetic decoding in the first step. The second step explored the variance in eidetic decoding attributable to NARA-3 rate and FastaReada scores. The results are presented in **Table 5**.

Age and non-verbal reasoning contributed significantly to the regression model,  $F(2,121) = 15.21$ ,  $p < 0.001$ , and accounted for 20.1% of the variation in eidetic decoding. The addition of NARA-3 reading rate and FastaReada variables explained an additional 49.6% of variation in eidetic decoding and this change in  $R^2$  was significant,  $F(2,119) = 97.37$ ,  $p < 0.001$ . FastaReada was the strongest unique predictor of visual word recognition strategy utilization (14.4%). Performance on the NARA-3 rate subtest uniquely predicted 2.6% of the variance in visual word recognition strategy utilization. Together, the five independent variables accounted for 69.7% of the variance in eidetic decoding.

## DISCUSSION

The purpose of this study was to evaluate if FastaReada can be utilized as a valid measure of reading skills. The relationship between FastaReada performance and comprehension was also examined, as was the importance of visual recognition in comparison to phonological skills for fast and accurate reading. The results are discussed in relation to the three hypotheses.

The first hypothesis that FastaReada scores would be associated with scores from the Accuracy, Comprehension and Rate subtests from the NARA-3 was strongly supported. The rate of words read accurately per minute on FastaReada was well correlated with individual subtest results on the NARA-3. Children who scored higher on FastaReada had greater accuracy rates for the words in the assigned prose passages of the NARA-3 reader, greater degrees of understanding of each passage, and

were able to decode more words accurately per minute than those who scored lower on FastaReada. The strong and significant correlations obtained in this study between FastaReada and the three subtests on the well-established, well-validated, and reliable measure of reading skills, the NARA-3 (Moorehouse and Yule, 1974; Neale, 1999) indicate that FastaReada meets the test of criterion-related validity. FastaReada thus provides a valid and direct measure of accuracy and speed, the two major accepted components of reading fluency. The findings are also consistent with research proposing that reading speed is a key factor in reading comprehension (Perfetti and Hogaboam, 1975; Jenkins et al., 2003; Danne et al., 2005; Rasinski et al., 2005; Yovanoff et al., 2005). The current results support the notion that conscious attentional demands for decoding inhibit understanding (Shankweiler, 1999; Klauda and Guthrie, 2008; Cutting et al., 2009).

In accordance with the second hypothesis, the contribution of visual word recognition was significantly greater than that of phonetic decoding skills (16.7% compared to 1.7%). These results demonstrate that fluent readers have already acquired the skills to decode novel words with speed and accuracy, and that these skills are predominantly reliant on competence in visual word recognition. Certainly, it has been demonstrated in the current study that the increased utilization of a phonetic decoding strategy, when reading, is detrimental to reading rate. In line with these results, training in phonemic awareness has been shown to improve reading difficulties at the word-level, contributing to more accurate word identification (Torgesen et al., 2001; Hatcher et al., 2006; Bowyer-Crane et al., 2008). However, improvements in single word reading accuracy do not imply improvements in continuous text reading rate. Indeed, slow reading rate often remains into adulthood despite remediation of decoding skills (Torgesen et al., 2001; O'Connor et al., 2007; Laycock and Crewther, 2008). Torgesen et al. (2001) reasoned that a limited repertoire for visually recognizable words results in increased reliance on phonemic analysis or guessing from context for word identification, and that this is inversely related to reading rate.

Findings from the second hypothesis also showed that fast and accurate readers tend to have higher levels of PA and phonetic decoding skills. The results showed a small but significant contribution of phonetic decoding skill to fast and accurate reading (1.7%). However, the results did not support a unique role for PA in fast and accurate reading after controlling for the contributions of visual word recognition and phonetic decoding. PA has been shown to be an important predictor for future reading in preliterate children (Gallagher et al., 2000; Snowling et al., 2003; Puolakanaho et al., 2004). Yet, the predictive nature of PA for reading ability has been documented to decrease with maturation of reading skills. Wagner et al. (1997) examined the changing relationship between PA and reading ability in a large longitudinal study. They documented a decrease in the unique contribution of PA to reading from 23% to only 4% between kindergarten and the fourth year of schooling after the contributions of word reading skills and vocabulary were taken into account. This is likely to reflect a shift toward an increasing reliance on a visual recognition strategy for the decoding of words that have become increasingly familiar with years of practice

**TABLE 5 | Hierarchical regression results of non-verbal reasoning, age in years, reading rate scores and FastaReada scores predicting eidetic decoding.**

Predictor	Final summary			Step summary	
	Beta	r	Sr	$\Delta R^2$	p
Step 1				0.201	<0.001
Age (years)	0.31	0.35	0.30		
Non-verbal reasoning	0.28	0.33	0.28		
Step 2				0.496	<0.001
Age (years)	0.09	0.35	0.09		
Non-verbal reasoning	0.03	0.33	0.03		
Reading rate	0.24	0.73	0.16		
FastaReada	0.59	0.81	0.38		

$N = 124$ .

and the contribution of a growing vocabulary. Thomson et al. (2006) demonstrated similar findings in young learner readers. The findings from the current study reinforce the findings of Thomson et al. (2006).

The final hypothesis that visual word recognition would be more strongly associated with performance on FastaReada than on the NARA-3 was supported. FastaReada was found to be better able to tap into the reader's efficacy in visual word recognition than the NARA-3 Rate subtest. An important feature of FastaReada is its utilization of an adaptive staircase routine for stimulus presentation. The adaptive staircase algorithm allows FastaReada to determine the shortest exposure time necessary for accurate visual word recognition whilst ensuring reliability of the measure. Stimulus exposure time in FastaReada, which becomes shorter with each correct response, can become so brief with increasing reading skill levels, that the accurate verbalization of text at the time of presentation becomes unachievable due to the motor limits of verbalization. Readers are therefore forced to read silently in order to encode the text and then verbalize them after they disappear. This method allows FastaReada to tap into visual word recognition and memory for series of words as facilitated by discourse-based anticipation. Thus, FastaReada can provide a more accurate measure of reading speed than traditional reading measures as it is not constrained by motor limits associated with the verbalization of text. This feature is particularly pertinent as adult reading is usually performed silently (Miller and Smith, 1989; Kragler, 1995).

The current study has shown FastaReada to be a valid measure of reading fluency through its strong positive correlations with established NARA tests of reading speed and accuracy for connected comprehensible text. Additional support for its criterion validity as a reading fluency measure was obtained through its strong relationship with established measures of eidetic decoding ability and phonetic decoding skills, as well as its moderate to strong relationship with the NARA reading comprehension measure. FastaReada therefore has the potential to play a valuable role in the education system. As an assessment tool that does not require specialist training, FastaReada can be used by educators to screen baseline-reading abilities and to monitor progress in the development of reading skills. By providing an indication of student progress, FastaReada would allow educators to develop and provide more effective, individualized reading programs that promote and nurture the development of reading fluency. FastaReada results can also alert educators and clinicians to the need for further assessment for better identification of the individual's specific difficulty.

We acknowledge that the current study has associated limitations. Previous studies, including our own (Rutkowski et al., 2003; Cotton et al., 2005a; Alloway, 2006; Laycock et al., 2006; Thomson et al., 2006; Vidyasagar and Pammer, 2010) indicate the important role attention and working memory contribute to reading abilities as regulatory and mediating factors respectively. Thus once validation studies for FastaReada are complete further investigations attention and working memory as contributory variables to FastaReada performance will be beneficial. This future research will be important as FastaReada requires contributions of working memory

when stimulus presentation is so short that examinees cannot verbalize the text during the exposure time. Additionally, the current research would have benefitted from a longitudinal design for the examination of the predictive validity of FastaReada. A longitudinal study would have enhanced findings related to the contribution of eidetic and phonological based abilities to reading skills over the course of reading development. In addition to phonological skills, rapid automatized naming tasks have been shown to be one of the best predictors of reading fluency. Future validity studies for FastaReada would therefore benefit from the inclusion of rapid automatized naming tasks in their design. Clearly, the next step required in the development of FastaReada as a measure of reading fluency is the provision of normative data, the determination of appropriate cut-off points for each year level, and possibly the design of a range of appropriate prose passages for the different year levels and for test and retest conditions.

## CONCLUSION

The current study has tested the criterion-related validity of FastaReada, a brief, computer-generated test of reading fluency that does not require specialist training for administration. FastaReada demonstrated a valid measure of the core features of reading fluency, speed and accuracy, as demonstrated by strong correlations with the established measures of accuracy, rate and comprehension on the NARA. FastaReada performance was also strongly correlated with measures of eidetic and phonetic text decoding abilities that have often been associated with the development of reading skills. FastaReada therefore provides a means for educators, clinicians, and researchers to quickly obtain a measure of reading fluency with relative confidence. The use of such a tool could contribute to more effective individualized reading instruction and remediation through the monitoring of reading fluency development. The current study also drew attention to the rapid, automatic, and visual nature of fluent reading. Results showed that while reading rate and accuracy is associated with PA, PA is not a predictor of fluency when the contributions of visual word recognition and phonetic decoding skills are taken into account. The multifaceted nature of reading skills is becoming increasingly recognized within the literature. This study adds to the burgeoning literature on reading fluency that is branching away from the traditional focus on phonological skills and exploring deficits in areas such as attention, working memory, and visual recognition. Further evaluation of FastaReada is warranted to assess reliability and to determine the most appropriate cut-off points for children of different year levels.

## ACKNOWLEDGMENTS

The authors would like to sincerely thank Dr. Robin Laycock for providing technical assistance and proof reading and, Dr. Ben Ong for his assistance with data analysis, and Jessica Peters for assisting with data collection.

## REFERENCES

- Alloway, T. P. (2006). How does working memory work in the classroom? *Educ. Res. Rev.* 1, 134–139.
- Blaklock, K. E. (2004). The importance of letter knowledge in the relationship between phonological awareness and reading. *J. Res. Read.* 27, 36–57. doi: 10.1111/j.1467-9817.2004.00213.x
- Boder, E. (1973). Developmental dyslexia: a diagnostic approach based on three atypical reading-spelling patterns. *Dev. Med. Child Neurol.* 15, 663–687. doi: 10.1111/j.1469-8749.1973.tb05180.x
- Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., et al. (2008). Improving early language and literacy skills: differential effects of an oral language versus phonology with reading intervention. *J. Child Psychol. Psychiatry* 49, 422–432. doi: 10.1111/j.1469-7610.2007.01849.x
- Breznitz, Z. (1987). Increasing first graders' reading accuracy and comprehension by accelerating their reading rates. *J. Educ. Psychol.* 79, 236–242. doi: 10.1037/0022-0663.79.3.236
- Breznitz, Z., and Misra, M. (2003). Speed of processing of the visual–orthographic and auditory–phonological systems in adult dyslexics: the contribution of “asynchrony” to word recognition deficits. *Brain Lang.* 85, 486–502. doi: 10.1016/S0093-934X(03)00071-3
- Castles, A., and Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition* 47, 149–180. doi: 10.1016/0010-0277(93)90003-E
- Castles, A., and Nation, K. (eds). (2006). *How does orthographic learning happen?* New York, NY: Psychology Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cotton, S. M., Crewther, D. P., and Crewther, S. G. (2005a). Measurement error: implications for diagnosis and discrepancy models of developmental dyslexia. *Dyslexia* 11, 186–202. doi: 10.1002/dys.298
- Cotton, S. M., Kiely, P. M., Crewther, D. P., Thomson, B., Laycock, R., and Crewther, S. G. (2005b). A normative and reliability study for the Raven's Coloured Progressive Matrices for primary school aged children from Victoria, Australia. *Pers. Individ. Differ.* 39, 647–659. doi: 10.1016/j.paid.2005.02.015
- Cutting, L. E., Materek, A., Cole, C. A., Levine, T. M., and Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Ann. Dyslexia* 59, 34–54. doi: 10.1007/s11881-009-0022-0
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- Danne, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., and Oranje, A. (2005). *Fourth-Grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading*. The Nation's Report Card. NCES 2006-469. Washington, DC: National Center for Education Statistics.
- Del Campo, R., Buchanan, W. R., Abbott, R. D., and Berninger, V. W. (2015). Levels of phonology related to reading and writing in middle childhood. *Read. Writ.* 28, 183–198. doi: 10.1007/s11145-014-9520-5
- Deno, S. L. (2003). Developments in curriculum-based measurement. *J. Special Educ.* 37, 184–192. doi: 10.1177/00224669030370030801
- Ehri, L. C. (1992). “Reconceptualizing the development of sight word reading and its relationship to recoding,” in *Reading Acquisition*, eds P. B. Gough, L. C. Ehri, and R. Treiman (Hillsdale, NJ: Lawrence Erlbaum Associates), 107–144.
- Ehri, L. C., and Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Appl. Psycholinguist.* 1, 371–385. doi: 10.1017/S0142716400009802
- Ehri, L. C., and Wilce, L. S. (1985). Movement into reading: is the first stage of printed word learning visual or phonetic? *Read. Res. Q.* 20, 163–179. doi: 10.2307/747753
- Fuchs, L. S., and Deno, S. L. (1991). Effects of curriculum within curriculum-based measurement. *Except. Child.* 58, 232–243.
- Fuchs, L. S., and Fuchs, D. (2002). Curriculum-based measurement: describing competence, enhancing outcomes, evaluating treatment effects, and identifying treatment nonresponders. *Peabody J. Educ.* 77, 64–84. doi: 10.1207/S15327930PJE7702-6
- Fuchs, L. S., Fuchs, D., Hosp, M. K., and Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: a theoretical, empirical, and historical analysis. *Sci. Stud. Read.* 5, 239–256. doi: 10.1207/S1532799XSSR0503\_3
- Gallagher, A., Frith, U., and Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *J. Child Psychol. Psychiatry* 41, 203–213. doi: 10.1111/1469-7610.00601
- Griffin, J. R., and Walton, H. N. (1987). *Dyslexia Determination Test: Examiner's Instruction Manual*. Los Angeles: Instructional Materials and Equipment Distributors.
- Hatcher, P. J., Hulme, C., Miles, J., Carroll, J. M., Hatcher, J., Gibbs, S., et al. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: a randomised controlled trial. *J. Child Psychol. Psychiatry* 47, 820–827. doi: 10.1111/j.1469-7610.2005.01559.x
- Hecht, R., Crewther, D., and Crewther, S. (2004). Rate of learning and asymptotic performance in an automatization task and the relation to reading. *Percept. Mot. Skills* 99, 1103–1121. doi: 10.2466/pms.99.3f.1103-1121
- Hosp, J. L., and Suchey, N. (2014). Reading assessment: reading fluency, reading fluently, and comprehension - Commentary on the special topic. *Sch. Psychol. Rev.* 43, 59–68.
- Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., and Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *J. Educ. Psychol.* 95, 719–729. doi: 10.1037/0022-0663.95.4.719
- Kameenui, E. J., and Simmons, D. C. (2001). The DNA of reading fluency. *Sci. Stud. Read.* 5, 203–210. doi: 10.1207/S1532799XSSR0503\_1
- Kinsey, K., Rose, M., Hansen, P., Richardson, A., and Stein, J. (2004). Magnocellular mediated visual-spatial attention and reading ability. *Neuroreport* 15, 2215–2218. doi: 10.1097/00001756-200410050-00014
- Klauda, S. L., and Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *J. Educ. Psychol.* 100, 310–321. doi: 10.1037/0022-0663.100.2.310
- Kragler, S. (1995). The transition from oral to silent reading. *Reading Psychol.* 16, 395–408. doi: 10.1080/0270271950160402
- Kuhn, M. R., Schwanenflugel, P. J., and Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: automaticity, prosody, and definitions of fluency. *Read. Res. Q.* 45, 230–251. doi: 10.1598/RRQ.45.2.4
- La Berge, D., and Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. doi: 10.1016/0010-0285(74)90015-2
- Laycock, R., and Crewther, S. G. (2008). Towards an understanding of the role of the ‘magnocellular advantage’ in fluent reading. *Neurosci. Biobehav. Rev.* 32, 1494–1506. doi: 10.1016/j.neubiorev.2008.06.2002
- Laycock, R., Crewther, S. G., Kiely, P. M., and Crewther, D. P. (2006). Parietal function in good and poor readers. *Behav. Brain Funct.* 2:26. doi: 10.1186/1744-9081-2-26
- Liberman, I. Y. (1971). Basic research in speech and lateralization of language: some implications for reading disability. *Bull. Orton Soc.* 21, 71–87. doi: 10.1007/BF02663712
- Logan, G. D. (1997). Automaticity and reading: perspectives from the instance theory of automatization. *Read. Writ. Q.* 13, 123–146. doi: 10.1080/1057356970130203
- Meyer, M. S., and Felton, R. H. (1999). Evolution of fluency training: old approaches and new directions. *Ann. Dyslexia* 49, 283–306. doi: 10.1007/s11881-999-0027-8
- Miller, S. D., and Smith, D. E. (1989). Relations among oral reading, silent reading and listening comprehension of students at differing competency levels. *Lit. Res. Instr.* 29, 73–84. doi: 10.1080/19388079009558006
- Moorehouse, A. J., and Yule, W. (1974). A comparison of the Neale and the Daniels and Diack reading tests. *Reading* 8, 24–27. doi: 10.1111/j.1467-9345.1974.tb00475.x
- Neale, M. D. (1999). *Neale Analysis of Reading Ability Manual*. Melbourne, VIC: ACER Press.
- O'Connor, O. E., White, A., and Lee Swanson, H. (2007). Repeated reading versus continuous reading: influences on reading fluency and comprehension. *Except. Child.* 74, 31–46. doi: 10.1177/001440290707400102
- O'Connor, R. E., Swanson, H. L., and Geraghty, C. (2010). Improvement in reading rate under independent and difficult text levels: influences on word and comprehension skills. *J. Educ. Psychol.* 102, 1–19. doi: 10.1037/a0017488
- Perfetti, C. A., Bell, L. C., and Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: evidence from backward masking. *J. Mem. Lang.* 27, 59–70. doi: 10.1016/0749-596X(88)90048-4



- Perfetti, C. A., and Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *J. Educ. Psychol.* 67, 461–469. doi: 10.1037/h0077013
- Petscher, Y., and Kim, Y.-S. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *J. Sch. Psychol.* 49, 107–129. doi: 10.1016/j.jsp.2010.09.004
- Puolakanaho, A., Poikkeus, A. M., Ahonen, T., Tolvanen, A., and Lyytinen, H. (2004). Emerging phonological awareness differentiates children with and without familial risk for dyslexia after controlling for general language skills. *Ann. Dyslexia* 54, 221–243. doi: 10.1007/s11881-004-0011-2
- Rasinski, T. V., Padak, N. D., McKeon, C. A., Wilfong, L. G., Friedauer, J. A., and Heim, P. (2005). Is reading fluency a key for successful high school reading? *J. Adolesc. Adult Lit.* 49, 22–27. doi: 10.1598/JAAL.49.1.3
- Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Section 2: coloured progressive matrices introducing the parallel version of the test. Great Britain: Oxford Psychologist Press.
- Rutkowski, J. S., Crewther, D. P., and Crewther, S. G. (2003). Change detection is impaired in children with dyslexia. *J. Vis.* 3, 95–105. doi: 10.1167/3.1.10
- Samuels, S. J. (2006). "Reading fluency: Its past, present and future," in *Fluency Instruction: Research-Based Best Practices*, 2nd Edn, eds T. V. Rasinski, C. Blachowicz, and K. Lems (New York, NY: Guilford), 3–16.
- Sereno, S. C., and Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends Cogn. Sci.* 7, 489–493. doi: 10.1016/j.tics.2003.09.010
- Shankweiler, D. (1999). Words to meanings. *Sci. Stud. Read.* 3, 112–127. doi: 10.1207/s1532799xssr0302\_2
- Snowling, M. J., Gallagher, A., and Frith, U. (2003). Family risk of dyslexia is continuous: individual differences in the precursors of reading skill. *Child Dev.* 74, 358–373. doi: 10.1111/1467-8624.7402003
- Swanson, H. L., Zheng, X., and Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: a selective meta-analysis of the literature. *J. Learn. Disabil.* 42, 260–287. doi: 10.1177/0022219409331958
- Tabachnick, B. G., and Fidell, L. S. (2013). *Using Multivariate Statistics* 6th Edn. Boston, MA: Pearson Publications.
- Taft, M. (1986). Lexical access codes in visual and auditory word recognition. *Lang. Cogn. Process.* 1, 297–308. doi: 10.1080/01690968608404679
- Thomson, B., Crewther, D. P., and Crewther, S. G. (2006). Wots that werd? Pseudowords (non-words) may be a misleading measure of phonological skills in young learner readers. *Dyslexia* 12, 289–299. doi: 10.1002/dys.328
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., and Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities immediate and long-term outcomes from two instructional approaches. *J. Learn. Disabil.* 34, 33–58. doi: 10.1177/002221940103400104
- Treiman, R., and Zukowski, A. (2013). "Levels of phonological awareness," in *Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman*, eds S. A. Brady and D. P. Shankweiler (New York, NY: Routledge).
- Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., and Newman, H. (2010). Oral reading fluency assessment: issues of construct, criterion, and consequential validity. *Read. Res. Q.* 45, 270–291. doi: 10.1598/RRQ.45.3.1
- Vellutino, F. R. (1977). Alternative conceptualizations of dyslexia: evidence in support of a verbal-deficit hypothesis. *Harv. Educ. Rev.* 47, 334–354. doi: 10.17763/haer.47.3.u117j10167686115
- Vellutino, F. R. (1979). *Dyslexia: Theory and Research*. Cambridge, MA: MIT Press.
- Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., and Chen, R. S. (2007). Components of reading ability: multivariate evidence for a convergent skills model of reading development. *Sci. Stud. Read.* 11, 3–32. doi: 10.1080/10888430709336632
- 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
- 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: 10.1037/0033-2909.101.2.192
- Wagner, R. K., Torgesen, J. K., and Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: new evidence of bi-directional causality from latent variable longitudinal study. *Dev. Psychol.* 30, 73–87. doi: 10.1037/0012-1649.30.1.73
- Wagner, R. K., Torgesen, J. K., and Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing: Examiner's Manual*. Austin: PRO-ED.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing causal relations between phonological processing abilities and word-level reading as children develop from beginning to fluent readers: a five year longitudinal study. *Dev. Psychol.* 33, 468–479. doi: 10.1037/0012-1649.33.3.468
- Wechsler, D. (2007). *Wechsler Individual Achievement Test*, 2nd Edn. Sydney: Pearson Clinical and Talent Assessment.
- Wesson, M. D. (1993). Diagnosis and management of reading dysfunction for the primary care optometrist. *Optom. Vis. Sci.* 70, 357–368. doi: 10.1097/00006324-199305000-00004
- Yovanoff, P., Duesbery, L., Alonzo, J., and Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educ. Meas.* 24, 4–12. doi: 10.1111/j.1745-3992.2005.00014.x
- 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

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Elhassan, Crewther, Bavin and Crewther. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Discrete versus multiple word displays: a re-analysis of studies comparing dyslexic and typically developing children

Pierluigi Zoccolotti<sup>1,2\*</sup>, Maria De Luca<sup>2</sup> and Donatella Spinelli<sup>2,3</sup>

<sup>1</sup> Department of Psychology, Sapienza University of Rome, Rome, Italy, <sup>2</sup> Neuropsychology Unit, Istituto di Ricovero e Cura a Carattere Scientifico – Fondazione Santa Lucia, Rome, Italy, <sup>3</sup> Department of Human Movement Sciences and Health, University of Rome “Foro Italico”, Rome, Italy

## OPEN ACCESS

### Edited by:

Simone Aparecida Capellini,  
São Paulo State University, Brazil

### Reviewed by:

Michael S. Dempsey,  
Boston University Medical Center,  
USA

Shuyan Sun,  
University of Maryland,  
Baltimore County, USA

### \*Correspondence:

Pierluigi Zoccolotti,  
Department of Psychology, Sapienza  
University of Rome, Via dei Marsi 78,  
00176 Rome, Italy  
pierluigi.zoccolotti@uniroma1.it

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 11 June 2015

**Accepted:** 22 September 2015

**Published:** 07 October 2015

### Citation:

Zoccolotti P, De Luca M and Spinelli D  
(2015) Discrete versus multiple word  
displays: a re-analysis of studies  
comparing dyslexic and typically  
developing children.  
Front. Psychol. 6:1530.  
doi: 10.3389/fpsyg.2015.01530

The study examines whether impairments in reading a text can be explained by a deficit in word decoding or an additional deficit in the processes governing the integration of reading subcomponents (including eye movement programming and pronunciation) should also be postulated. We report a re-analysis of data from eleven previous experiments conducted in our lab where the reading performance on single, discrete word displays as well multiple displays (texts, and in few cases also word lists) was investigated in groups of dyslexic children and typically developing readers. The analysis focuses on measures of time and not accuracy. Across experiments, dyslexic children are slower and more variable than typically developing readers in reading texts as well as vocal reaction time (RTs) to singly presented words; the dis-homogeneity in variability between groups points to the inappropriateness of standard measures of size effect (such as Cohen's *d*), and suggests the use of the ratio between groups' performance. The mean ratio for text reading is 1.95 across experiments. Mean ratio for vocal RTs for singly presented words is considerably smaller (1.52). Furthermore, this latter value is probably an overestimation as considering total reading times (i.e., a measure including also the pronunciation component) considerably reduces the group difference in vocal RTs (1.19 according to Martelli et al., 2014). The ratio difference between single and multiple displays does not depend upon the presence of a semantic context in the case of texts as large ratios are also observed with lists of unrelated words (though studies testing this aspect were few). We conclude that, if care is taken in using appropriate comparisons, the deficit in reading texts or lists of words is appreciably greater than that revealed with discrete word presentations. Thus, reading multiple stimuli present a specific, additional challenge to dyslexic children indicating that models of reading should incorporate this aspect.

**Keywords:** reading, dyslexia, text reading, multiple displays, vocal reaction time

## Introduction

Reading a passage is a complex task requiring a number of sub-componential tasks, which start from the perception of visual features (contours, segments of various orientations), then letters, and word recognition to continue with the integration of successive words into a coherent stream. At this level, syntactic and semantic processing allows for the identification of the sentence meaning and the possibility to place it within the more general context of the text. All this takes place in association with motor processing, i.e., saccades and fixations to scan the text, and pronunciation. In reading deficiencies, it is interesting to understand which is the level of analysis which is most appropriate to describe the reading difficulty (here, we restrict our analysis to developmental deficits, i.e., developmental dyslexia DD). Potentially, any of the above listed levels may generate the difficulty as research has clearly shown they are all necessary steps in the reading process. So, one may think that the deficit in DD originates as early in the information processing chain as in the elaboration of letters; alternatively, one may see the deficit originating at a word locus or later when the identification of several words is merged as it occurs in the reading of meaningful texts. Note that early deficits (including also motor processing such as eye movements) may spread into later processing as a cascade effect. As an example, if we imagine a child to be impaired in letter recognition (or in the programming of eye movements) this will severely affect all subsequent processing, including word recognition, integration of decoding, and pronunciation etc.

So, one very general question, which has been extensively examined in the literature on DD, is which is the earliest level of processing at which a deficit can be reliably found. It is generally held that children with dyslexia are spared in processing letters. Importantly, evidence is based on a variety of sensitive techniques (such as contrast thresholds, or masked tachistoscopic presentation) that guarantee that this sparing is not due to the lack of sensitivity of the measures used (Bosse et al., 2007; Lassus-Sangosse et al., 2008; Martelli et al., 2009; De Luca et al., 2010). By contrast, it is well established that children with dyslexia are selectively impaired in processing strings of letters (whether forming existing words or not). Indeed, major models of reading (such as the dual route cascaded model or DRC; Coltheart et al., 2001; the CDP+ model; Perry et al., 2007; and the triangle model, Plaut et al., 1996) are focused in explaining reading at the word level. So, up to date evidence indicates that the nuclear deficit in DD is at the level of letter orthographic string decoding.

However, there is reason to think that the reading deficit may not be entirely explained at the word level and that the need to integrate the processing of words with other sub-components of reading may represent an additional burden, which selectively affects the reading of dyslexic children. So, a second general question is whether impairments at subsequent levels of processing can be identified and explained either as independent defects or due to a cascade effect from deficits in orthographic decoding.

Critical to answer this question is the comparison between single, discrete word displays (typical of experimental settings) and multiple displays (as it occurs in the reading of meaningful texts). However, comparing such different levels of processing may prove difficult, *in primis* due to variations in general difficulty of the two tasks. A further difficulty is that different measures are typically used. When single words are examined a frequently used measure is vocal reaction time (RT), i.e., the time between the stimulus onset and the beginning of subject's vocal response. When texts or lists of words are examined the reading time also includes the time required to utter the sentences (or the words in the list).

Therefore, examining total reading times (i.e., RTs plus pronunciation times) also in the case of singly presented words may be instrumental to compare reading fluency between discrete and multiple displays. In one such study, we observed that typically developing readers showed an advantage on multiple with respect to discrete items: they were able to process the next stimulus while uttering the current word indicating that pronunciation times overlapped with decoding times (Zoccolotti et al., 2013). By contrast, children with dyslexia did not show the advantage for multiple over discrete stimuli in the case of lists of short words and actually showed a disadvantage in the case of long words (on which they were slower than in the case of discrete stimuli). We proposed that the disproportionate impairment of children with dyslexia in dealing with multiple arrays indicates a difficulty in integrating the multiple subcomponents of the reading task over and above the basic nuclear deficit in decoding words (Zoccolotti et al., 2013).

Can we re-evaluate the previous literature in light of the findings indicating a specific deficit in reading sub-components in dyslexia? The main question of the present study is whether impairments in functional reading can be explained by the basic deficit in letter string decoding or an additional deficit in the integration of various reading subcomponent should also be postulated. To this aim we report a re-analysis of data from previous experiments conducted in our laboratory where the reading performance on both single, discrete word and multiple words has been investigated in groups of typically developing and dyslexic children. The analysis focuses on measures of time and not accuracy.

Our first question is whether the reading deficit shown by children with dyslexia is greater with discrete or multiple visual displays. Clearly, the experimental conditions used in our previous studies are not ideal for this comparison. On the one hand, studies on single words typically reported RTs not reading times (i.e., a measure including pronunciation, as in Zoccolotti et al., 2013); thus, one should ideally control for the effect of pronunciation on the results of previous studies. On the other hand, single and multiple stimuli were not matched in terms of stimulus characteristics. Studies based on single word presentation usually aimed to understand the effect on vocal RTs of parameters such as word frequency, word length, morphological structure and so on, often leading to a large number of levels of the experimental manipulations. By contrast, multiple word displays were texts or list of words; these materials

are typically used to select the groups of dyslexic and typically developing children according to their basic reading skills and often yield a single measure of overall performance. Thus, to compare the efficiency in reading words in multiple and single stimulus displays we have to average data collected over different experimental conditions in discrete word studies to obtain an overall estimate of the reading time also for singly presented words.

Additional methodological questions arise in the case of such comparison. Namely, which is the appropriate index to compare the size of the difference between dyslexic and control readers? How can the difference in dependent measure (RTs versus total reading times) be controlled for? Does the presence of a meaningful context modulate the performance of children with dyslexia? The way we tackle each of these questions is detailed below along with the presentation of results.

## Methods

### Selection Criteria of Target Studies

We focused on studies in which children with dyslexia were compared to a group of typically developing readers using very similar (although not identical) subject's selection criteria. We also limited the analysis to groups of children attending sixth grade, which was the most common age in our previous studies. With these criteria we were able to trace eleven different studies where we had both measures of text reading (used for the purpose of screening by standard reading text) and measures of vocal RTs to single word (used for the specific aims of the given study). All but one recently completed study have been previously published. Some of these studies also included different screening tests requiring the reading of lists of words (see below for more details).

### Reading Measures

The basic reading test used for screening purposes was the *MT Reading Test* (Cornoldi and Colpo, 1995): a passage adapted for children's age is presented and the child is requested to read it as fast and accurately as possible. Two tests requiring the reading of lists of words were used. One was the *Words and Non-words Reading Test* (Zoccolotti et al., 2005). This features four lists of 30 words varying for frequency and length; separate norms are available for each of the four sub-lists. Another test was the word sub-test from the *Battery for the Evaluation of Developmental Dyslexia and Dysgraphia* (Sartori et al., 1995). A total of 112 words are presented in four 28-word sub-lists varying for frequency and imageability. However, only a single measure is usually reported for this test as norms report only this measure of general performance. In both tests, the list of words was printed vertically; the task, as in the *MT Reading Test* and vocal RTs, was to read the words as fast and accurately as possible.

Reaction times were measured in all studies by presenting a word in the center of a computer screen; the word was visible until the children started his/her uttering. The RT was measured as the interval between stimulus onset and vocal onset.

## Results

### Fluency Differences in Reading Texts

**Table 1** presents the list of studies selected, indicating the number of dyslexic children and chronologically matched typically developing children considered in each of them. A total of 331 typically developing children and 172 cases participated to the studies. The mean times for reading a standard text passage (*MT Reading Test*; Cornoldi and Colpo, 1995) are reported for both groups. The mean reading times are expressed in terms of s per word (by averaging words of different length in the paragraph). Various observations can be advanced based on the data in the table.

As expected, children with dyslexia have higher mean reading times than typically developing readers. On average, their reading times (1.05 s per word) are about twice as slow as those of typically developing children (0.54 s per word). Thus, there is a mean 1.95 ratio between the performance of the two groups (the range of ratios across studies is 1.4–2.4).

Second, dyslexic children are also considerably more variable in their performance. Mean SD is 0.37 in dyslexic children and only 0.11 in typically developing children. Thus, there is co-variance between mean performance and variability, a finding often reported in the RT literature (Wagenmakers and Brown, 2007). Notably, the larger inter-individual variability shown by children with dyslexia goes beyond the proportionality between mean and SD. This is shown by the coefficient of variation values (i.e., the ratio between SD and mean). In all studies the coefficient of variation for dyslexic children is higher than that of typically developing children (mean value = 0.34 for dyslexic children and 0.19 for typically developing children). This finding underscores the difficulty in comparing the two groups through standard parametric analyses. Indeed, these data indicate a strong and systematic violation of the homogeneity assumption, which is critical to apply parametric analyses. These observations are supported by comparisons through the Levene test for equality of variances. In all studies, the test indicated that the variances of the two groups were significantly different (at least,  $p < 0.01$ ).

This large difference in variability points to the inappropriateness of using standard measures of size effects, such as Cohen's  $d$  or  $\eta^2$  which assume homogeneity of variance. In computing  $d$ , one can use the SD from either sample (as they are assumed to be homogeneous; Cohen, 1988) or, possibly, the mean of the two. However, results would drastically and systematically change if the SD of either group is used. For example, if one computes the Cohen's  $d$  value on the first study in **Table 1** (Judica et al., 2002), one obtains very different values depending on which standard deviation is used to calculate  $d$ . It is 3.64 by using the SD of typically developing children (0.25), 1.10 using the SD of dyslexic children (0.82), and 1.69 using the average between the two SDs. While all these values indicate a large effect it is clear that the estimate of effect size depends heavily upon which SD value is used. In conclusion, standard effect sizes (such as  $d$  or  $\eta^2$ ) do not appear to capture the main effect of reading deficiency. This is better described as a multiplicative effect. As such, a better descriptor of the effect is provided by the ratio that

**TABLE 1 |** Reading times at the MT Reading test.

Reference	Typically developing children				Children with dyslexia				Groups' performance ratio
	<i>N</i>	Mean	<i>SD</i>	Coefficient of variation	<i>N</i>	Mean	<i>SD</i>	Coefficient of variation	
Judica et al., 2002	16	0.62	0.25	0.40	18	1.51	0.82	0.54	2.4
Spinelli et al., 2005	23	0.56	0.14	0.25	17	1.10	0.42	0.38	2.0
Barca et al., 2006	68	0.55	0.10	0.18	14	0.94	0.26	0.28	1.7
Zoccolotti et al., 2006	12	0.52	0.09	0.18	10	1.10	0.48	0.43	2.1
Burani et al., 2008	34	0.56	0.09	0.16	17	1.02	0.32	0.31	1.8
De Luca et al., 2008	34	0.56	0.09	0.16	17	1.02	0.32	0.31	1.8
Zoccolotti et al., 2008	29	0.55	0.12	0.22	6	1.31	0.61	0.47	2.4
Paizi et al., 2011	36	0.50	0.09	0.19	18	0.87	0.21	0.24	1.8
Paizi et al., 2013	17	0.48	0.05	0.10	17	0.86	0.19	0.22	1.8
Martelli et al., 2014	43	0.52	0.09	0.17	25	1.14	0.38	0.33	2.2
Unpublished data	19	0.50	0.06	0.12	13	0.71	0.13	0.18	1.4
<b>Total</b>	<b>331</b>				<b>172</b>				
<b>Mean</b>		<b>0.54</b>	<b>0.11</b>	<b>0.19</b>		<b>1.05</b>	<b>0.37</b>	<b>0.34</b>	<b>1.95</b>

Mean reading times (seconds per word read) and SDs are reported for typically developing children and children with dyslexia from the listed studies; the size of each group is also reported. The last column reports the dyslexics/controls means ratios.

captures the multiplicative nature of the performance difference between dyslexic and control readers. Clearly, samples from the various studies show different performances. However, the ratios between the performances between the two groups are relatively stable across studies, ranging from 1.4 to 2.4 with an average close to 2.

## Comments

All parametric analyses rest on the homogeneity of variance assumption. Thus, researchers are typically reluctant in abandoning such a basic tenet. A number of data transformations are often adopted to approach normality of distribution and to control for as much as possible of dis-homogeneities of variance. One such example is the log-transformation often used with RTs. In the case of text reading, sometimes time measures (such as s per word) are converted to speed measures (word per s; for a discussion of the advantages and limits of this transformation see Toraldo and Lorusso, 2012). In this perspective, deviations from normality and from homogeneity of variance are seen as accidental perturbations in the data set that need to be corrected for. In contrast, large inter-individual variability is typically associated to developmental/learning phases, and the huge inter-individual variability in DD is an expression of their condition of being still in a early learning phase of reading, whereas at the same age typically developing readers have reached a plateau in their reading performance.

Present data suggest an interesting alternative to the solution of correcting for deviation from homogeneity of variance. Variabilities between the groups are actually truly dis-homogenous as impaired reading is systematically associated to increased individual variability. The prediction of increased SD in DD stems quite clearly from models that aim to account for the presence of global components in the data. For example, within the rate and amount model (RAM) Faust et al. (1999) propose that, when the difference between two

groups is accounted for by a global factor, one expects means of different conditions to linearly covariate with the SDs of the corresponding conditions. Further comments on this perspective will be advanced in the section “Group differences in reading: Linear-additive versus multiplicative models” of the Discussion. Throughout the study we will use the ratio between groups’ performance as an index that capture the multiplicative nature of the performance difference between dyslexic and typically developing readers.

## Fluency Differences in Reading Discrete Words

**Table 2A** reports data on single word reading derived from the same studies as in **Table 1**. Mean vocal RTs are reported. Note that different studies used different experimental manipulations, such as length, frequency, morphological structure etc. However, due to our current interest, we report here both data for single conditions and averaged data across conditions.

An inspection of the table indicates a number of relevant findings. Clearly, children with dyslexia are slower than typically developing children across conditions. All studies in this re-analysis, showed a highly significant main effect of the group factor (with at least  $p < 0.01$ ) at standard Anovas. However, the ratios between the two groups are consistently lower than those in **Table 1**. Across studies and experimental manipulations the overall mean ratio is 1.52 (range across experiments from 1.28 to 1.89; range across all experimental manipulations 1.12 and 2.13); this mean value is considerably lower than that for text reading (1.95, see **Table 1**). Thus, the slowing of dyslexic children with respect to typically developing readers is about 95% in text reading and only 52% in the case of single word reading. If, instead of averaged data, we separately compare the between groups’ ratios for each of the 69 experimental conditions, in only two cases are the ratios above the mean value (1.95) obtained for text reading (see **Table 2A**).



**TABLE 2A | Vocal reaction times (RTs) in several experimental conditions from the listed studies.**

	Typically developing children			Children with dyslexia			Groups' performance ratio
	Mean	SD	Coefficient of variation	Mean	SD	Coefficient of variation	
Judica et al., 2002							
2-letter word	559	57	0.10	742	147	0.20	1.33
3-letter word	574	67	0.12	856	285	0.33	1.49
4-letter word	585	67	0.11	960	433	0.45	1.64
5-letter word	633	87	0.14	1103	535	0.48	1.74
Mean	588			915			1.55
Spinelli et al., 2005							
3-letter word	639	90	0.14	844	274	0.32	1.32
4-letter word	646	111	0.17	982	479	0.49	1.52
5-letter word	646	107	0.16	1035	435	0.42	1.60
6-letter word	687	191	0.28	1247	565	0.45	1.81
7-letter word	758	173	0.23	1477	657	0.44	1.95
8-letter word	744	209	0.28	1412	643	0.45	1.90
Mean	686			1166			1.68
Barca et al., 2006							
HF words with contextual rules	685	104	0.15	850	95	0.11	1.24
HF words without contextual rules	675	105	0.16	862	92	0.11	1.28
LF words with contextual rules	768	126	0.16	1000	153	0.15	1.30
LF words without contextual rules	725	127	0.18	955	134	0.14	1.32
Mean	713			917			1.28
Zoccolotti et al., 2006							
4-letter word	708	93	0.13	997	207	0.21	1.41
5-letter word	716	105	0.15	1088	222	0.20	1.52
6-letter word	745	137	0.18	1306	392	0.30	1.75
7-letter word	784	138	0.17	1371	523	0.38	1.75
Mean	738			1190			1.61
Burani et al., 2008							
Derived words (Exp2)	672	114	0.17	1086	485	0.45	1.62
Simple words	670	108	0.16	1103	487	0.44	1.65
Mean	671			1095			1.63
De Luca et al., 2008							
4-letter word	612	63	0.10	816	246	0.30	1.33
5-letter word	624	74	0.12	886	294	0.33	1.42
6-letter word	622	70	0.11	912	305	0.33	1.47
7-letter word	649	90	0.14	1039	441	0.42	1.60
8-letter word	691	103	0.15	1110	480	0.43	1.61
9-letter word	710	104	0.15	1153	496	0.43	1.62
Mean	651			986			1.51
Zoccolotti et al., 2008							
4-letter word	601	67	0.11	851	297	0.35	1.42
5-letter word	612	65	0.11	803	221	0.27	1.31
6-letter word	599	72	0.12	983	420	0.43	1.64
7-letter word	655	82	0.12	1012	332	0.33	1.55
8-letter word	611	82	0.13	1047	641	0.61	1.71
Mean	616			939			1.53
Paizi et al., 2011							
HF 4-letter word - pure list	540	57	0.10	685	137	0.20	1.27
HF 5-letter word - pure list	553	62	0.11	713	127	0.18	1.29
HF 6-letter word - pure list	566	62	0.11	779	165	0.21	1.38
HF 7-letter word - pure list	558	69	0.12	783	165	0.21	1.40

(Continued)

TABLE 2A | Continued

	Typically developing children			Children with dyslexia			Groups' performance ratio
	Mean	SD	Coefficient of variation	Mean	SD	Coefficient of variation	
LF 4-letter word - pure list	555	60	0.11	732	134	0.18	1.32
LF 5-letter word - pure list	568	69	0.12	807	165	0.20	1.42
LF 6-letter word - pure list	579	69	0.12	843	205	0.24	1.46
LF 7-letter word - pure list	595	83	0.14	974	328	0.34	1.64
HF 4-letter word - mixed list	563	57	0.10	702	111	0.16	1.25
HF 5-letter word - mixed list	582	62	0.11	786	174	0.22	1.35
HF 6-letter word - mixed list	601	74	0.12	842	199	0.23	1.40
HF 7-letter word - mixed list	590	72	0.12	817	174	0.21	1.39
LF 4-letter word - mixed list	587	64	0.11	753	149	0.20	1.28
LF 5-letter word - mixed list	598	67	0.11	798	139	0.17	1.33
LF 6-letter word - mixed list	610	79	0.13	876	222	0.25	1.44
LF 7-letter word - mixed list	626	88	0.14	956	265	0.28	1.53
<i>Mean</i>	580			803			1.38
<b>Paizi et al., 2013</b>							
HF word (Exp.1)	597	50	0.08	794	72	0.09	1.33
LF word (Exp.1)	632	53	0.08	898	150	0.17	1.42
HF 4-letter word (Exp.3)	510	78	0.15	679	98	0.14	1.33
HF 5-letter word (Exp.3)	534	74	0.14	741	106	0.14	1.39
HF 6-letter word (Exp.3)	543	86	0.16	812	212	0.26	1.49
HF 7-letter word (Exp.3)	539	78	0.14	808	167	0.21	1.50
LF 4-letter word (Exp.3)	544	80	0.15	753	102	0.13	1.38
LF 5-letter word (Exp.3)	568	88	0.15	833	166	0.20	1.47
LF 6-letter word (Exp.3)	568	97	0.17	864	192	0.22	1.52
LF 7-letter word (Exp.3)	580	94	0.16	959	311	0.32	1.65
<i>Mean</i>	562			814			1.45
<b>Martelli et al., 2014</b>							
4-letter word	506	59	0.12	827	361	0.44	1.63
5-letter word	511	69	0.13	939	433	0.46	1.84
6-letter word	536	76	0.14	1056	502	0.47	1.97
7-letter word	541	95	0.17	1150	585	0.51	2.13
<i>Mean</i>	524			993			1.89
<b>Unpublished data</b>							
HF 4-letter word	490	40	0.08	547	58	0.10	1.12
HF 5-letter word	480	40	0.08	568	72	0.13	1.18
HF 6-letter word	491	45	0.09	594	84	0.14	1.21
HF 7-letter word	497	43	0.09	589	88	0.15	1.18
LF 4-letter word	496	42	0.08	564	56	0.10	1.14
LF 5-letter word	507	38	0.07	594	66	0.11	1.17
LF 6-letter word	509	40	0.08	611	77	0.13	1.20
LF 7-letter word	514	50	0.10	613	67	0.11	1.19
<i>Mean</i>	498			585			1.17
<b>Total mean*</b>	<b>602.5</b>	<b>82.6</b>	<b>0.13</b>	<b>891.7</b>	<b>265.3</b>	<b>0.28</b>	<b>1.47</b>
<b>SD</b>	<b>75.2</b>			<b>202.9</b>			<b>1.52</b>

Mean vocal RTs (milliseconds per word), standard deviations and coefficients of variation for typically developing children and children with dyslexia are reported. The last column reports the dyslexics/controls mean ratios. \*Total means do not include partial means.

Notably, values vary across experimental manipulations. In particular, in studies manipulating length (as in the first one by Judica et al., 2002) there is a clear tendency for ratios to increase as a function of stimulus length (in this case from 1.33 to 1.74 with progressively longer words). The same is apparent in most

(Spinelli et al., 2005; Zoccolotti et al., 2006, 2008; De Luca et al., 2008; Paizi et al., 2011, 2013; Martelli et al., 2014) although not all (research with unpublished data) studies. The other variable that has been manipulated most often is frequency. Across all contrasts between high and low frequency words, the ratios

between the two groups for the high frequency words averaged 1.31 while those for the low frequency words averaged 1.38. Thus, ratios do not vary appreciably between conditions as a function of frequency. It may be interesting to compare these findings to the calculations based on more sophisticated methods, such as the analyses based on the RAM by Faust et al. (1999) which were carried out in several of the quoted studies. This may help understanding the efficacy, and limits, of the procedure of using the ratio as an estimate of size of the group differences in reading skills; further comments on this question will be proposed in the Discussion section.

Dyslexic children are considerably more variable as a group than typically developing children; their average SD is 265.3 ms while that of control readers is only 82.6 ms. In general, variability grows as a function of the general difficulty of the experimental conditions with more difficult conditions yielding larger SD. Across conditions there is a 0.81 correlation ( $p < 0.001$ ) between means and SDs in control children; the correlation is 0.86 ( $p < 0.001$ ) for dyslexic children. These results are in keeping with the general law indicating a relationship between condition means and standard deviations for RT measures (Wagenmakers and Brown, 2007). Furthermore, also coefficients of variation are about twice as high in dyslexic children (mean value = 0.28) than in control readers (mean value = 0.13); for only 6 out of 69 conditions were the coefficients of variation higher for control than dyslexic readers. Comparisons with the Levene test indicated that the variances of the two groups were significantly different (with at least  $p < 0.05$ ) in 53 out of 69 comparisons.

## Comments

Despite variations across studies and experimental conditions, the ratio data clearly indicate that vocal RTs of dyslexic children are slower than typically developing readers by about 50%. This contrasts with the ratios measured for reading times,

where dyslexic children were about 100% slower than typically developing children.

## From RTs to Total Reading Times in Reading Discrete Words

One general finding of the above analyses is that ratios between the performance of dyslexic and typically developing readers in the case of multiple stimulus displays are higher than in the case of discrete stimulus displays. Clearly, the two sets of data refer to different measures. In the case of discrete stimuli only the time between stimulus presentation and the *incipit* of the response is considered but not the actual pronunciation time. By contrast, in the case of multiple stimulus displays the measure is the total reading time (i.e., it includes pronunciation time). So, one may consider how the use of different measures affect the results.

A way to tackle this problem is to include pronunciation time measures in experiments with single stimulus displays. RTs and pronunciation times together give a measure of total reading time, which may be usefully compared to the mean reading time per item in the case of multiple stimulus displays. Measuring pronunciation times is simple although time consuming as it requires trial-by-trial analysis. A few studies have used this procedure in recent times (e.g., Davies et al., 2013). One of the studies in **Table 2A** also adopted this procedure (Martelli et al., 2014).

Martelli et al.'s (2014) results for pronunciation times and total reading times (i.e., RTs plus pronunciation times) are presented in **Table 2B** and can be compared with RT data for the same study presented in the low part of **Table 2A**. Across conditions the ratio between the performance of dyslexic and control readers is 1.89 for RTs in this particular study (i.e., a value in the high range compared to similar studies in the same table). The ratio (see **Table 2B**) is close to unity in the case of pronunciation times (1.04); thus, across conditions children

**TABLE 2B | Pronunciation times and total reading times from Martelli et al.'s (2014) study.**

	Typically developing children			Children with dyslexia			Groups' performance ratio
	Mean	SD	Coefficient of variation	Mean	SD	Coefficient of variation	
Pronunciation time							
4-letter word	395	60	0.15	404	51	0.13	1.02
5-letter word	461	59	0.13	468	69	0.15	1.02
6-letter word	487	63	0.13	510	71	0.14	1.05
7-letter word	564	75	0.13	596	83	0.14	1.06
Mean	477	64	0.14	494	68.5	0.14	1.04
Total reading time							
4-letter word	901	89	0.10	1231	366	0.30	1.37
5-letter word	972	103	0.10	1406	455	0.32	1.45
6-letter word	1023	101	0.10	1565	522	0.33	1.53
7-letter word	1106	127	0.11	1746	611	0.35	1.58
Mean	1000	105	0.10	1487	489	0.33	1.48

Mean pronunciation time and mean total reading time (milliseconds per single word), standard deviations and coefficients of variation are reported for typically developing children and children with dyslexia. The last column reports the dyslexics/controls mean ratio.

with dyslexia show pronunciation times very similar to those of control children and also very similar inter-individual variability (as indicated by both SDs and coefficients of variations). When considering total reading times, the groups' performance ratio is 1.48, i.e., intermediate between those obtained with the two measures contributing to total reading time (i.e., vocal RTs and pronunciation). In particular, this value is much smaller than the one obtained in the same study in the case of RTs (1.89; see **Table 2A**).

We can use the values measured in this study to estimate the average drop of the mean ratio in the case of total reading time as opposed to vocal RTs to discrete words. The proportion

$$1.52 : x = 1.89 : 1.48$$

where 1.52 is the mean ratio for RTs across studies; and the two remaining values are the ratios for RTs and total reading time in the Martelli et al.'s (2014), study, respectively. The proportion leads to an estimated groups' ratio of 1.19 when total reading time of single words is considered. As this is based on a single study this is clearly a rough estimate of the groups' ratio for discrete word presentation. However, it generally indicates that the difference in groups' ratios between multiple (1.95) and discrete (1.52) displays is likely underestimated by the use of RTs rather than total reading times and is presumably much larger.

## Comments

Overall, the results indicate that the RT groups' ratios are presumably a high estimate of the groups' differences in single word reading, as RTs are only the part of the response that is most sensitive to the experimental manipulations. If one includes also the component of pronunciation, which distinguishes minimally between the two groups, the ratios drop substantially indicating that the differences in groups' ratios between multiple and discrete displays are much larger than those estimated based on overall text reading on the one side and vocal RTs to words (as in **Table 2**) on the other. Indeed, the present computations indicate a group ratio of 1.95 in the case of multiple displays (see **Table 1**) and an overall estimate of 1.19 in the case of discrete displays (according to the formula above); this is a quite large difference in size effect. If confirmed by subsequent studies (it would be interesting that future studies also consider total reading times in RTs experiments), this pattern of findings would indicate that efficiency in reading aloud single words plays only a moderate role in determining the fluency of dyslexic children when reading texts, which would certainly be a surprising finding. An important, and generally neglected role would be played by the other components involved in the reading task.

## Reading Lists of Words

One additional confounding factor when comparing reading texts with reading isolated words is the presence of contextual information only in the former, but not the latter, case. So, the larger group differences in text reading may depend upon a selective difficulty of dyslexic children to integrate the semantic context. Indeed, there is reason to consider this hypothesis

unlikely. Children with dyslexia do not show a selective deficit in comprehending texts at least in the case in which no time limit is imposed, as in the standard procedure of the *MT Reading Test* (Cornoldi and Colpo, 1995). Typically, under these conditions, dyslexic children show only a mild defect or even an entirely spared performance (e.g., Zoccolotti et al., 1999). Still, one could envisage the hypothesis that, at least under conditions in which both speed and accuracy are encouraged (as it is required to the children in the standard *MT Reading Test*), the need for an ongoing integration of successive pieces of information may provide an additional burden widening the performance difference between the two groups.

Information on this question may come from conditions in which the child is asked to read lists of unrelated words printed on a page. Under these conditions, no role of context is present and no need to integrate the meaning of successive information is required for effective performance. In some of the studies listed in **Table 1** we also used two such tests (*Words and Non-words Reading Test*; Zoccolotti et al., 2005, and the word list from the *Battery for the Evaluation of DD and Dysgraphia*; Sartori et al., 1995). In the former test four separate measures are taken for words varying for frequency and length; for the latter test a single measure is usually reported (based on available norms).

For four studies, there are data on the *Words and Non-words Reading Test* (see **Table 3A**). Across studies and conditions there is a ratio of 1.83 (range 1.51–2.26) between the performance of dyslexic and control readers. This estimate is lower than the one observed in the case of text reading (1.95) but higher than the one for single word reading (1.52) particularly if one considers the need for a correction due to the use of RTs rather than total reading times. On average, the ratios are slightly higher for low (1.89) than high (1.77) frequency words, and higher for long (1.96) than short (1.71) words. As in previous comparisons, dyslexic children were more variable than typically developing children, both in terms of SDs (0.74 vs. 0.24, respectively) and of coefficients of variation (0.47 vs. 0.27, respectively). Comparisons with the Levene test indicated that the variances of the two groups were significantly different (with at least  $p < 0.05$ ) in 14 out of 16 comparisons.

As to the word list from the *Battery for the Evaluation of DD and Dysgraphia* (Sartori et al., 1995) there are data available from two of the studies with information on discrete word reading (see **Table 3B**). In all studies the ratio between the performance of dyslexic and control readers was above 2 (mean = 2.62), a value higher than that in the case of text reading.

## Comments

The data available in the case of word lists are fewer than those on text reading and the results are also somewhat scattered with higher ratios for the word list from the *Battery for the Evaluation of DD and Dysgraphia* (Sartori et al., 1995) than for the word lists from the *Words and Non-words Reading Test* (Zoccolotti et al., 2005). Differences in list composition probably account for this effect although it is at present difficult to understand which feature in the list composition is critical



**TABLE 3A | Reading times (s per word) for the Words and Non-words Reading Test (Zoccolotti et al., 2005) from the listed studies.**

	Typically developing children			Children with dyslexia			Groups' performance ratio	
	Mean	SD	Coefficient of variation	Mean	SD	Coefficient of variation		
Spinelli et al., 2005								
4-to-5-letter high frequency words	0.76	0.21	0.28	0.97	0.50	0.52	1.27	
8-to-10-letter high frequency words	0.99	0.25	0.25	1.62	0.78	0.48	1.63	
4-to-5-letter low frequency words	0.86	0.25	0.29	1.25	0.73	0.59	1.46	
8-to-10-letter low frequency words	1.38	0.47	0.34	2.30	0.90	0.39	1.66	
Mean	1.00			1.53				1.51
Zoccolotti et al., 2006								
4-to-5-letter high frequency words	0.61	0.18	0.29	0.98	0.43	0.44	1.61	
8-to-10-letter high frequency words	0.78	0.20	0.25	1.61	0.82	0.51	2.06	
4-to-5-letter low frequency words	0.69	0.19	0.28	1.25	0.58	0.46	1.82	
8-to-10-letter low frequency words	1.09	0.31	0.28	2.31	0.96	0.41	2.13	
Mean	0.79			1.54				1.91
Burani et al., 2008								
4-to-5-letter high frequency words	0.68	0.20	0.29	1.04	0.41	0.40	1.53	
8-to-10-letter high frequency words	0.87	0.22	0.25	1.45	0.62	0.43	1.67	
4-to-5-letter low frequency words	0.76	0.22	0.28	1.33	0.48	0.36	1.74	
8-to-10-letter low frequency words	1.21	0.34	0.28	2.03	0.61	0.30	1.68	
Mean	0.88			1.46				1.66
Zoccolotti et al., 2008								
4-to-5-letter high frequency words	0.62	0.13	0.21	1.21	0.76	0.62	1.96	
8-to-10-letter high frequency words	0.80	0.15	0.19	1.96	1.08	0.55	2.43	
4-to-5-letter low frequency words	0.73	0.17	0.23	1.66	1.10	0.66	2.29	
8-to-10-letter low frequency words	1.18	0.34	0.29	2.82	1.15	0.41	2.38	
Mean	0.83			1.91				2.26
Total mean*	0.87	0.24	0.27	1.61	0.74	0.47	1.83	1.83
SD	0.23			0.54				

Means, standard deviations, and coefficients of variation for typically developing children and children with dyslexia are reported. The last column reports the dyslexics/controls mean ratio. \*Total means do not include partial means.

**TABLE 3B | Reading times (s per word) for the word list from the Battery for the Evaluation of Developmental Dyslexia (DD) and Dysgraphia (Sartori et al., 1995) from the listed studies.**

	Typically developing children			Children with dyslexia			Groups' performance ratio
	Mean	SD	Coefficient of variation	Mean	SD	Coefficient of variation	
Judica et al., 2002	0.74	0.30	0.41	1.81	0.81	0.45	2.44
Spinelli et al., 2005	0.74	0.30	0.41	2.07	0.90	0.43	2.80
<b>Mean</b>	<b>0.74</b>	<b>0.30</b>	<b>0.41</b>	<b>1.94</b>	<b>0.86</b>	<b>0.44</b>	<b>2.62</b>
<b>SD</b>	<b>0.00</b>			<b>0.19</b>			

Means, standard deviations, and coefficients of variation for typically developing children and children with dyslexia are reported. The last column reports the dyslexics/controls mean ratio.

to yield such outcome. However, data from word lists are generally in keeping with the idea that reading multiple words generates greater group differences than reading discrete words. This occurs in the absence of any contextual effect. Thus, it appears that the requirement to read a sequence of stimuli rather than a single one is sufficient to generate a large size group difference also in the absence of a meaningful semantic context.

It should be added that these data do not allow excluding the possibility that the context exerts some at least partial effect in modulating the group differences in reading fluency. To obtain a definite response on this point would require stimuli which vary only along the context dimension; e.g., comparing regular and scrambled matched texts may be instrumental to clarify this question. In this respect, it should be noted that the possible direction of such an effect is not obvious. On the one

side, one could envisage that, since they have generally spared semantic skills, dyslexic children may actually be favored by the presence of contextual information. On the other, one could hypothesize that in a time demanding task the need to online process the information concerning the syntactic relationship between words represents an additional burden, which further dampens performance. *Ad hoc* research is needed to clarify this point. However, the present data seem sufficiently clear to indicate that the need to process multiple stimuli poses by itself a selective stress on dyslexic children such that their difference in performance with control readers becomes much more pronounced than that observed in the case of discrete displays.

## Discussion

Comparing the performance of dyslexic and typically developing readers in tasks such as reading texts, lists of words and single words poses challenging methodological questions and the present data only represent an initial sketch of the complex set of relationships that may influence reading fluency. Furthermore, it seems important that the present data should be supported by additional evidence from other research groups. However, even the available evidence seems strong enough to conclude that, at least for Italian language, reading multiple stimuli present a specific challenge to the dyslexic children at the sixth grade of schooling indicating that models of reading should incorporate this aspect (e.g., Zoccolotti et al., 2014). By contrast, up to date most models of reading are based on the assumption that the reading process can be explained at the single word level (Plaut et al., 1996; Coltheart et al., 2001; Perry et al., 2007).

The reviewed data seem sufficiently persuasive to conclude that group differences (dyslexic vs. typically developing readers) in reading fluency in the case of multiple word displays are much greater than differences in the case of discrete word displays. In fact, as shown above, the difference between the two sets of data are presumably larger than they appear based on available data. In the case of discrete stimulus presentations, typically RTs are presented; this measure extracts the portion of the response that is most sensitive to the decoding differences. However, if one considers a measure (total reading time) that is more similar to that used in text or words list reading, a much greater difference emerges between discrete and multiple displays.

### Deficits in Multiple Displays

Clear differences in reading isolated words are present between typically developing and dyslexic readers. However, dyslexic readers have larger deficits compared to typically developing readers when they have to deal with multiple displays. Reading in these conditions requires integration of various sub-components. While processing the ongoing word, the reader has to perform some parafoveal analysis of the next word, to program the more effective landing of the next forward saccade (often skipping functional words; for a review see Rayner, 2009). The output

of word processing is held in memory in order to effectively synchronize the pronunciation of the stimulus with the decoding of the subsequent words (referred to as eye-voice lead; Fairbanks, 1937). Reading under these conditions selectively dampens dyslexic performance. Thus, it appears that, in understanding the reading impairment of dyslexic children, one has also to explain this failure with multiple stimuli and not limit the interpretation to the deficit at a single-word level.

Why should dyslexic children be selectively impaired in dealing with multiple visual displays? One can envisage four possible scenarios.

Firstly, one could consider the text reading deficit as a cascade effect of the nuclear defect in orthographic decoding. The deficit might be amplified through the greater complexity, and henceforth difficulty, involved in text reading. According to this view, even if the reading deficits for discrete and multiple displays have different sizes (the latter being greater than the former), they would essentially refer to the same deranged mechanism. Within this hypothesis, the deficit with discrete displays should accurately predict the one with multiple displays. By contrast, there is evidence that, in accounting for individual differences in text reading fluency, the performance on rapid automatized naming (RAN) tasks (Denckla and Rudel, 1974) increases the variance explained by single word reading in Greek (Protopapas et al., 2013) and Italian (Zoccolotti et al., 2014) readers. This finding is not in keeping with the idea that a single deficit explains impairments with discrete and multiple displays.

Second, it is conceivable that, in addition to the decoding deficit (which is clearly evident also in the present re-analysis), dyslexic children have a selective deficit in one of the other reading subcomponents. While it is likely that at least some of the children may have additional defects, previous attempts along this line have been generally unsuccessful. For example, as shown above, articulation deficits are absent (e.g., Martelli et al., 2014). As for a deficit in the programming and execution of eye movements as suggested in an early study (Pavlidis, 1981), most successive evidence has been inconsistent with this hypothesis (e.g., Brown et al., 1983; Olson et al., 1983; De Luca et al., 1999); i.e., dyslexic children have eye movements comparable to controls except when dealing with reading material. Further, in spite of their deranged pattern of eye movements during reading, impaired readers show an intact mechanism for performing corrective re-fixations (a mechanism linked to oculomotor and visual processes not linguistic ones; Gagl et al., 2014). Although some researchers are still working on the hypothesis that some selective deficits in eye movements programming or execution may actually be impaired in dyslexic children (e.g., Bucci et al., 2008) this hypothesis seems poorly supported by evidence. Overall, the available results do not seem strong enough to account for the large differences in text reading fluency although it is difficult to reach definite conclusions on this literature.

A third scenario is to focus on the possible interaction of the various sub-components underlying multiple word reading with the reading deficit. Even though none of the sub-components (apart from orthographic decoding) reveals a selective deficit

(as envisaged in the case of the second scenario), the presence of a deficit in orthographic decoding could make the multi-task management considerably more difficult (De Luca et al., 2013). For example, in this view, dyslexic children would not be impaired in parafoveal processing *per se*. However, the need to process the next (right) word parafoveally to appropriately calibrate the successive saccade may be hindered by the attention of the child being fully focused on the ongoing target word in the troubled attempt to process it. There is some evidence supporting this view (Yan et al., 2013). Overall, one could posit that a set of processes, which are in themselves spared, represent an attention overload due to the presence of a selective deficit in orthographic decoding. In this interactive view, orthographic decoding would indirectly dampen text reading fluency as it may prove difficult to carry out a complex task if one does not manage well one of the task sub-components (De Luca et al., 2013). This third scenario does not require any additional deficit (as in the second scenario) or amplification (as in the “cascade” first scenario) other than the defect in orthographic decoding. However, one may imagine that factors, such as divided attention, may interact with the decoding deficit in modulating the reading fluency of children with dyslexia. According to this interactive view (and differently from the cascade view) one would not expect the single-word decoding deficit to accurately predict the deficit with multiple words. Furthermore, one would not expect performance on divided attention tasks and/or executive tasks to directly correlate with reading. However, one could put forward the hypothesis that performance on these tasks may act as suppressor variables allowing for increased prediction in the case of reading words in multiple (but not single) displays. Communality analyses may allow the detection of such suppression effects. Overall, integrating several sub-components of the reading task may pose an additional, partially independent, challenge to the dyslexic children (Zoccolotti et al., 2014).

A fourth scenario to explain the greater fluency deficit of dyslexic children with multiple than single word displays focuses on the difference between the experimental conditions used in the two sets of tasks. In the single condition, the word is abruptly displayed on the screen; in the multiple conditions, the words are statically displayed on a sheet of paper (or a PC screen; the support does not probably make a critical difference). It is well known that the abrupt onset of a stimulus is perceptually salient, captures bottom-up attention (Jonides and Yantis, 1988), elicits prepotent and fast saccades (McDowell et al., 2008), and triggers fast visual processing up to target identification (indicated by shorter RTs in search tasks; e.g., Theeuwes, 1994) or word decoding (indicated by reading rate increment in Rapid Serial Visual Presentation task; Rubin and Turano, 1992). By contrast, reading in the static condition of a multiple display implies a more internally driven visual scanning of the items; saccades (and decoding) are self-paced and driven by parafoveal pre-analysis (Schotter et al., 2012). It is likely that these differences between static and dynamic reading conditions are relevant for the overall speed of processing. Indeed, the neural network involved in self-paced and externally triggered movements do not entirely overlap and have different time constants (Thickbroom and Mastaglia,

1985; Cunnington et al., 2002). Consistently, some recent EEG (Dimigen et al., 2011, 2012) and fMRI (Choi et al., 2014; Richlan et al., 2014) studies investigating the neural basis of reading have privileged the ecological method of sentence reading rather than single-word reading or rapid serial visual presentation. In this perspective, single word presentation may facilitate reading processing by automatic recruitment of attention and by providing an external pacing of the reading activity; this facilitation might be particularly advantageous (in terms of speed) for dyslexic children with respect to typically developing readers. Some authors described the “sluggish” attention (Hari and Renvall, 2001) of dyslexic children. This defect would be partially overcome by abrupt presentation of stimuli. In other terms, an externally triggered onset of the target word would make the reading of dyslexic children more “automatized”, that is, more similar to the reading of typically developing readers. Consequently, the difference between groups would be less marked in the case of single stimulus displays. The very high correlation between text reading and individual speed in RAN tasks (where multiple color patches or objects have to be named) but not in single color naming (when the color patch is abruptly displayed on the screen) may be seen as supporting this line of interpretation (Georgiou et al., 2013). To test this hypothesis, it may prove instrumental to compare reading of multiple word displays in conditions in which the observer is requested to read words at his/her own self-pace or some external abrupt cue (such as a bar underlining the target word) introduces an imperative stimulus in the display. Reading under externally paced conditions is expected to yield smaller group differences between dyslexic and control readers.

The present evidence is still too sparse to definitively choose among these alternatives. However, some facts seem clear. In particular, the lack of a strong correlation between performance on discrete and multiple displays (de Jong, 2011) is inconsistent with the first “cascade” scenario. Also, the search for selective deficits in eye movements programming and execution has yet proven unsuccessful making also the second scenario unlikely. However, the last two scenarios seem promising venues for future research; some possible hypotheses worth testing have been outlined.

### Group Differences in Reading: Linear-Additive versus Multiplicative Models

As compared to typically developing children, dyslexic readers are not only much slower but also considerably more variable in their performance. This is indicated by much greater SDs and coefficients of variation. Thus, greater variability goes even beyond what might be anticipated on the basis of an increase in the mean performance. Multiplicative models may account for this pattern more effectively than linear additive models.

One such model is the RAM proposed by Faust et al. (1999). Accordingly, performance depends multiplicatively by an individual factor (the rate at which the individual processes information) and by a task related factor (the difficulty of the given experimental condition referred to as “amount”). Along this reasoning, performance on a given condition does not

merely express the specific ability to deal with a given specific condition but also depends upon more general factors such as the global ability of the individual to process information and the general difficulty of the task (over and above the specificity of the experimental condition). Note that this perspective generally indicates a situation often referred to as “task impurity”: i.e., there is a lot more in the performance in any given task than the specific process which is intended to probe. To express the rate factor in DD we have referred to a global factor in orthographic pre-lexical processing. In this view, individuals have a characteristic speed in processing orthographic materials which influences all conditions and tasks which require to visually process orthographic strings of letters. So, this factor is global in the sense that is not condition-specific but it affects all conditions within the orthographic domain (such as naming long and short words, high and low frequency words, naming non-words, lexical decision). However, it is not to be intended as “general” as it does not apply to task in which other types of stimuli are to be processed (e.g., naming objects; Zoccolotti et al., 2008) or word stimuli are to be processed in a sensory modality different from the visual one (i.e., with auditory presentation; Marinelli et al., 2011).

The present analyses indicate that the multiplicative nature of the difference between dyslexic and typically developing readers is well captured by ratios while it is not well accounted for by effect size measures (such as Cohen's  $d$ ) within the parametric linear additive perspective. In this context, ratios present advantages but also limitations. The main advantage is that they allow to quickly compare performances in otherwise disparate conditions, which would be difficult to compare within the rather selective requirements of models which aim to account for individual differences in performance in timely tasks. Here, we showed that the ratios for reading performance in the case of multiple visual displays are considerably higher than those for reading performance in the case of single visual displays.

An important limitation of using the ratio is that this value indicates an overall relationship between the performances of the two groups. By contrast, an attempt of models such as the RAM, the DEM or the diffusion model is to distinguish between different components of the response. So, according to Myerson et al. (2003) one could separate a decisional and a non-decisional part of the response (and clear predictions are put forward to tease out these two components of the response). Based on these predictions, Martelli et al. (2014) showed that it was only the decisional component of the response which contributed in generating the group differences in performance. Using a lexical decision, a similar conclusion was reached by Zeguers et al. (2011) who, based on a diffusion model analysis, observed no difference between dyslexic and control readers in the non-decision components of the RTs. Indeed, the diffusion model makes a step ahead and, beyond the distinction between decisional and non-decisional components, is also able to account for the possible modulating role of criterion (or “conservatoriness”) in mediating the group differences (Ratcliff, 1978). However, also in this case, experimental conditions are constrained within rather strict

requirements and it is not immediately apparent how group differences in multiple versus single stimulus displays could be examined within the experimental requirements envisaged by these models.

Empirically, it may be instructive to examine whether ratios capture effects in ways which are more or less compatible with the more tuned analyses performed in relationship with the above mentioned models. To test the possible presence of selective effects over and beyond the effect of the global factor in orthographic processing in a number of studies we referred to the RAM (Faust et al., 1999). This proposes a number of data transformations (including an individually based  $z$ -score transformation) which allow obtaining condition measures stripped off the effect of the global factor<sup>1</sup>. This transformation allows distinguishing between group by condition interactions which can be entirely ascribed to an over-additivity effect and those in which a residual, selective effect of a specific experimental variable is detectable. In several experiments we found that, if one examines raw RT data, dyslexic children show larger frequency effects than control readers. However, if one controls for the effect of the global factor by normalizing data over individual subjects as suggested by Faust et al. (1999), the group by frequency interaction disappears (Paizi et al., 2013). By contrast, in a number of studies we found that the effect of stimulus length was detected even after accounting for the effect of the global factor (e.g. Zoccolotti et al., 2008).

When we re-examine the results of these experiments by using ratios it is clear that frequency plays no detectable role in the case of RT studies (see Table 2) and a very limited role in the case of total reading times (Table 3A). By contrast, length exerts a very clear impact on ratio values in the case of RT data (Table 2) and some influence also in the case of total reading times (Table 3A). Therefore, it appears that, although results in terms of ratios represent less sophisticated measures of group differences than those that may be obtained with reference to models such as the RAM or DEM they yield a pattern of results which is generally consistent with that obtained with reference to these models. This reinforces the idea that the large difference in ratios between discrete and multiple displays is a genuine phenomenon, not one derived from the adoption of such a measure.

In conclusion, the idea that group differences in reading do not easily fit with linear additive models has indeed widespread implications. Nearly all the literature on reading skills uses parametric analyses based on linear additive assumptions. When deviations from normality are detected, appropriate data transformations (such as log transform in the case of RTs or speed, as opposed to time, measures in the case of texts) are used. Furthermore, it is generally held that results from ANOVAs are generally quite robust, in that they are not very sensitive

<sup>1</sup>The formal limits of using a ratio (or proportion) transformation have been discussed by Faust et al. (1999). Essentially, as this transformation identifies an overall relationship between two measures the results would be identical to those of using transformations, such as the  $z$  score or regression transformations, only in the case in which the additive constant (i.e., the intercept) of the relationship is null.



to deviations from normality. So, we certainly do not wish to claim that all results in the literature are faulty or unreliable. Rather, we would like to make the general point that it seems unfounded to try to explain by means of linear additive models differences which are clearly multiplicative. If seen within a linear additive model group differences are prone to be sensitive to over-additivity effects, i.e., more difficult conditions will generate larger group differences over and above the influence of a specific experimental manipulation. By contrast, examining the group differences from the perspective of multiplicative models (such as the RAM) may potentially allow separating the different factors that contribute in generating individual differences in performance.

## Conclusion

Children with dyslexia show a clear impairment in reading words when they are singly presented (and vocal RTs are measured). In particular, they are both slower and considerably more variable than typically developing readers. This pattern of results is consistent with the idea that the deficit is best expressed in terms of a multiplicative rather than additive difference. Thus, an effective way to describe the group difference is with the use

of ratios rather than standard measures of size effects (such as Cohen's *d*).

The RT measure is very sensitive to capture the part of the response most sensitive to the reading deficit. Thus, the very clear results obtained measuring RTs to single word presentation may give the impression that the reading deficit is strong and independent of the number of targets present in the display. However, if care is taken in using appropriate comparisons, it is clear that the deficit in reading texts or lists of words is appreciably greater than that revealed with discrete stimulus presentations. Thus, to fully explain the reading deficit of these children one should also account for their difficulty in managing the complex set of sub-component tasks underlying the fluent read a text. While several hypotheses can be put forward to explain this deficit, the present re-analysis underscores that an exhaustive explanation of the reading deficit cannot be obtained based on the performance on single word presentations only.

## Acknowledgments

This work was supported by grants from the Department of Health and Sapienza University of Rome.

## References

- \*Barca, L., Burani, C., Di Filippo, G., and Zoccolotti, P. (2006). Italian developmental dyslexic and proficient readers: where are the differences? *Brain Lang.* 98, 347–351. doi: 10.1016/j.bandl.2006.05.001
- 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
- Brown, B., Haegerstrom-Portnoy, G., Adams, A. J., Yingling, C. D., Galin, D., Herron, J., et al. (1983). Predictive eye-movements do not discriminate between dyslexics and control children. *Neuropsychologia* 21, 121–128. doi: 10.1016/0028-3932(83)90078-7
- Bucci, M. P., Breimond-Gignac, D., and Kapoula, Z. (2008). Latency of saccades and vergence eye movements in dyslexic children. *Exp. Brain Res.* 188, 1–12. doi: 10.1007/s00221-008-1345-5
- \*Burani, C., Marcolini, S., De Luca, M., and Zoccolotti, P. (2008). Morpheme-based reading aloud: evidence from dyslexic and skilled Italian readers. *Cognition* 108, 243–262. doi: 10.1016/j.cognition.2007.12.010
- Choi, W., Desai, R. H., and Henderson, J. M. (2014). The neural substrates of natural reading: a comparison of normal and nonword text using eyetracking and fMRI. *Front. Hum. Neurosci.* 8:1024. doi: 10.3389/fnhum.2014.01024
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Erlbaum.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. C. (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
- Cornoldi, C., and Colpo, G. (1995). *Nuove Prove di Lettura MT Per la Scuola Media Inferiore, Manuale (New MT Reading Test for the Middle School)*. Firenze: Organizzazioni Speciali.
- Cunnington, R., Windischberger, C., Deecke, L., and Moser, E. (2002). The preparation and execution of self-initiated and externally-triggered movement: a study of event-related fMRI. *Neuroimage* 15, 373–385. doi: 10.1006/nimg.2001.0976
- Davies, R., Rodriguez-Ferreiro, J., Suairez, P., and Cueto, 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. (2011). What discrete and serial rapid automatized naming can reveal about reading. *Sci. Stud. Read.* 5, 314–337. doi: 10.1080/10888438.2010.485624
- \*De Luca, M., Barca, L., Burani, C., and Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cogn. Behav. Neurol.* 21, 227–235. doi: 10.1097/WNN.0b013e318190d162
- De Luca, M., Burani, C., Paizi, D., Spinelli, D., and Zoccolotti, P. (2010). Letter and letter-string processing in developmental dyslexia. *Cortex* 46, 1272–1283. doi: 10.1016/j.cortex.2009.06.007
- De Luca, M., Di Pace, E., Judica, A., Spinelli, D., and Zoccolotti, P. (1999). Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia* 37, 1407–1420. doi: 10.1016/S0028-3932(99)00038-X
- De Luca, M., Pontillo, M., Primativo, S., Spinelli, D., and Zoccolotti, P. (2013). The eye-voice lead during oral reading in developmental dyslexia. *Front. Hum. Neurosci.* 7:696. doi: 10.3389/fnhum.2013.00696
- Denckla, M. B., and Rudel, R. G. (1974). Rapid “automatized” naming of pictured objects, colors, letters and numbers by normal children. *Cortex* 10, 186–202. doi: 10.1016/S0010-9452(74)80009-2
- Dimigen, O., Kliegl, R., and Sommer, W. (2012). Trans-saccadic parafoveal preview benefits in fluent reading: a study with fixation-related brain potentials. *Neuroimage* 62, 381–393. doi: 10.1016/j.neuroimage.2012.04.006
- Dimigen, O., Sommer, W., Hohlfield, A., Jacobs, A. M., and Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. *J. Exp. Psychol. Gen.* 140, 552–572. doi: 10.1037/a0023885
- Fairbanks, G. (1937). The relation between eye-movements and voice in the oral reading of good and poor silent readers. *Psychol. Monogr.* 48, 78–107. doi: 10.1037/h0093394
- Faust, M. E., Balota, D. A., Spieler, D. H., and Ferraro, F. R. (1999). Individual differences in information processing rate and amount: implications for group differences in response latency. *Psychol. Bull.* 125, 777–799. doi: 10.1037/0033-2909.125.6.777
- Gagl, B., Hawelka, S., and Hutzler, F. (2014). A similar correction mechanism in slow and fluent readers after suboptimal landing positions. *Front. Hum. Neurosci.* 8:355. doi: 10.3389/fnhum.2014.00355

\*Studies included in the re-analysis.

- Georgiou, G. K., Parrila, R., Cui, Y., and Papadopoulos, T. C. (2013). Why is rapid automatized naming related to reading? *J. Exp. Child Psychol.* 115, 218–225. doi: 10.1016/j.jecp.2012.10.015
- 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
- Jonides, J., and Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Percept. Psychophys.* 43, 346–354. doi: 10.3758/BF03208805
- \*Judica, A., De Luca, M., Spinelli, D., and Zoccolotti, P. (2002). Training of developmental surface dyslexia modifies performance and eye fixation duration in reading. *Neuropsychol. Rehabil.* 12, 177–197. doi: 10.1080/09602010244000002
- Lassus-Sangosse, D., N'Guyen-Morel, M. A., and Valdois, S. (2008). Sequential or simultaneous visual processing deficit in developmental dyslexia? *Vision Res.* 48, 979–988. doi: 10.1016/j.visres.2008.01.025
- Marinelli, C. V., Angelelli, P., Di Filippo, G., and Zoccolotti, P. (2011). Is developmental dyslexia modality specific? A visual-acoustic comparison on Italian dyslexics. *Neuropsychologia* 49, 1718–1729. doi: 10.1016/j.neuropsychologia.2011.02.050
- \*Martelli, M., De Luca, M., Lami, L., Pizzoli, C., Pontillo, M., Spinelli, D., et al. (2014). Bridging the gap between different measures of the reading speed deficit in developmental dyslexia. *Exp. Brain Res.* 232, 237–252. doi: 10.1007/s00221-013-3735-6
- Martelli, M., Di Filippo, G., Spinelli, D., and Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *J. Vis.* 9, 1–18. doi: 10.1167/9.4.14
- McDowell, J. E., Dyckman, K. A., Benjamin Austin, B. P., and Clementz, B. A. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades: evidence from studies of humans. *Brain Cogn.* 68, 255–270. doi: 10.1016/j.bandc.2008.08.016
- Myerson, J., Hale, S., Zheng, Y., Jenkins, L., and Widaman, K. F. (2003). The difference engine: a model of diversity in speeded cognition. *Psychon. Bull. Rev.* 10, 262–288. doi: 10.3758/BF03196491
- Olson, R. K., Kliegl, R., and Davidson, B. J. (1983). Dyslexic and normal readers' eye movements. *J. Exp. Psychol. Hum.* 9, 816–825. doi: 10.1037/0096-1523.9.5.816
- \*Paizi, D., De Luca, M., Burani, C., and Zoccolotti, P. (2011). List context manipulation reveals orthographic deficits in Italian readers with developmental dyslexia. *Child Neuropsychol.* 17, 459–482. doi: 10.1080/09297049.2010.551187
- \*Paizi, D., De Luca, M., Zoccolotti, P., and Burani, C. (2013). A comprehensive evaluation of lexical reading in Italian developmental dyslexics. *J. Res. Read.* 36, 303–329. doi: 10.1111/j.1467-9817.2011.01504.x
- Pavlidis, G. T. H. (1981). Do eye movements hold the key to dyslexia? *Neuropsychologia* 19, 57–64. doi: 10.1016/0028-3932(81)90044-0
- Perry, C., Ziegler, J. C., and Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: the CDP<sup>+</sup> model of reading aloud. *Psychol. Rev.* 114, 273–315. doi: 10.1037/0033-295X.114.2.273
- 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. doi: 10.1037/0033-295X.103.1.56
- Protopapas, A., Altani, A., and Georgiou, G. K. (2013). Development of serial processing in reading and rapid naming. *J. Exp. Child Psychol.* 116, 914–929. doi: 10.1016/j.jecp.2013.08.004
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* 62, 1457–1506. doi: 10.1080/17470210902816461
- Richlan, F., Gagl, B., Hawelka, S., Braun, M., Schurz, M., Kronbichler, M., et al. (2014). Fixation-related fMRI analysis in the domain of reading research: using self-paced eye movements as markers for hemodynamic brain responses during visual letter string processing. *Cereb. Cortex* 24, 2647–2656. doi: 10.1093/cercor/bht117
- Rubin, G. S., and Turano, K. (1992). Reading without saccadic eye movements. *Vision Res.* 32, 895–902. doi: 10.1016/0042-6989(92)90032-E
- Sartori, G., Job, R., and Tressoldi, P. E. (1995). *Batteria Per la Valutazione Della Dislessia e Della Disortografia Evolutiva. (Battery for the Evaluation of Developmental Dyslexia and Dysgraphia)*. Firenze: Organizzazioni Speciali.
- Schotter, E. R., Angele, B., and Rayner, K. (2012). Parafoveal processing in reading. *Atten. Percept. Psychophys.* 74, 5–35. doi: 10.3758/s13414-011-0219-2
- \*Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., and Zoccolotti, P. (2005). Length effect in word naming latencies: role of reading experience and reading deficit. *Dev. Neuropsychol.* 27, 217–235. doi: 10.1207/s15326942dn2702\_2
- Theeuwes, J. (1994). Stimulus-driven capture and attentional set: selective search for color and visual abrupt onsets. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 799–806. doi: 10.1037/0096-1523.20.4.799
- Thickbroom, G. W., and Mastaglia, F. L. (1985). Cerebral events preceding self-paced and visually triggered saccades. A study of presaccadic potentials, Electroencephalogr. *Clin. Neurophysiol.* 62, 277–289. doi: 10.1016/0168-5597(85)90005-X
- Torraldo, A., and Lorusso, M. L. (2012). Syllables per second versus seconds per syllable when measuring reading speed. *Front. Psychol.* 16:518. doi: 10.3389/fpsyg.2012.00518
- Wagenmakers, E. J., and Brown, S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychol. Rev.* 114, 830–841. doi: 10.1037/0033-295X.114.3.830
- Yan, M., Pan, J., Laubrock, J., Kliegl, R., and Shu, H. (2013). Parafoveal processing efficiency in rapid automatized naming: a comparison between Chinese normal and dyslexic children. *J. Exp. Child Psychol.* 115, 579–589. doi: 10.1016/j.jecp.2013.01.007
- Zeguers, M. H., Snellings, P., Tijms, J., Weeda, W. D., Tamboer, P., Bexkens, A., et al. (2011). Specifying theories of developmental dyslexia: a diffusion model analysis of word recognition. *Dev. Sci.* 14, 1340–1354. doi: 10.1111/j.1467-7687.2011.01091.x
- Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., and Spinelli, D. (2005). *Prova di Lettura di Parole e Non Parole. (Words and Non-words Reading Test)*. Roma: IRCCS Fondazione Santa Lucia. Available at: <http://www.hsantalucia.it/modules.php?name=Content&pa=showpage&pid=1032>
- Zoccolotti, P., De Luca, M., Di Pace, E., Judica, A., Orlandi, M., and Spinelli, D. (1999). Markers of developmental surface dyslexia in a language (Italian) with high grapheme-phoneme correspondence. *Appl. Psycholing.* 20, 191–216.
- \*Zoccolotti, P., De Luca, M., Judica, A., and Burani, C. (2006). Delayed naming eradicates the word length effect in developmental dyslexia. *Brain Lang.* 99, 27–28. doi: 10.1016/j.bandl.2006.06.028
- \*Zoccolotti, P., De Luca, M., Judica, A., and Spinelli, D. (2008). Isolating global and specific factors in developmental dyslexia: a study based on the rate and amount model (RAM). *Exp. Brain Res.* 186, 551–560. doi: 10.1007/s00221-007-1257-9
- Zoccolotti, P., De Luca, M., Lami, L., Pizzoli, C., Pontillo, M., and Spinelli, D. (2013). Multiple stimulus presentation yields larger deficits in children with developmental dyslexia: a study with reading and RAN-type tasks. *Child Neuropsychol.* 19, 639–647. doi: 10.1080/09297049.2012.718325
- Zoccolotti, P., De Luca, M., Marinelli, C. V., and Spinelli, D. (2014). Modeling individual differences in text reading fluency: a different pattern of predictors for typically developing and dyslexic readers. *Front. Psychol.* 5:1374. doi: 10.3389/fpsyg.2014.01374

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Zoccolotti, De Luca and Spinelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

