12-9-2016

Subphonemic Sensitivity in Low Literacy Adults

Monica Y.C. Li
Psych Dept UConn, monica.li@uconn.edu

Subphonemic Sensitivity in Low Literacy Adults

Monica Yin-Chen Li

M.S., National Central University, Taiwan, 2013

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2016

Approval Page

Masters of Science Thesis

Subphonemic Sensitivity in Low Literacy Adults

Presented by

Monica Yin-Chen Li, M.S.

Major Advisor____

James S. Magnuson

Associate Advisor____

Gerry T.M. Altmann

Associate Advisor____

David Braze

Associate Advisor____

Whitney Tabor

University of Connecticut

2016

Acknowledgements

First and foremost, I would like to thank my major advisor, Dr. Jim Magnuson, for his guidance, encouragement, and patience over the past two and a half years, when I often felt lost and doubtful about my work and myself. I would also like to thank my committee member and collaborator, Dr. Dave Braze, for frequently making the time to discuss this project in depth with me, and for the financial support of my research assistantship from his grant. In addition, I would like to thank my other two committee members, Dr. Gerry Altmann and Dr. Whit Tabor, for their concrete comments and challenging questions throughout my writing process, which have made the revision process more enjoyable and helped me to look at this project from various perspectives more critically. I would also like to thank Haskins Laboratories for facilitating the collaboration regarding this project. Finally, I would like to thank my lab mates, close friends, and family for their intellectual and emotional support. Special thanks to my husband, Tim, for all the things that he has done to cheer me up and ease off my pressure during my most stressful moments.

Table of Contents

Abstract

The link between phonological abilities and reading skills has been well-established in both typical and atypical language development. However, the nature of the phonological deficits in poor readers remains a debated topic. While poor readers have been mostly assumed to have underspecified or "fuzzy" phonological representations (Tallal et al., 1998), the opposite alternative, *over*-specified phonological representations, has also been hypothesized (Serniclaes, 2006). To examine the two phonological hypotheses, the current study used the eye-tracking paradigm in the study of Dahan et al. (2001) to investigate individuals' sensitivity to subphonemic information in young adults with a wide range of reading abilities. Our findings suggested a trend of higher sensitivity to subphonemic information in lower-ability readers, consistent with the over-specification hypothesis. In addition, our sample with a lower range of socio-economic status highlighted the need to take environmental factors into consideration for theoretical and practical purposes in reading acquisition.

Introduction

Phonological deficits have been implicated in reading difficulties in both typical and atypical reading acquisition. Poor readers are often assumed to have underspecified or somehow “fuzzy” phonological representations (Elbro, 1998; Tallal, Merzenich, Miller, & Jenkins, 1998), although *over*-specified representations have also been hypothesized (Bogliotti, Serniclaes, Messaoud-Galusi, & Sprenger-Charolles, 2008; Serniclaes, 2006). This may be surprising, given well-established weaknesses by individuals with dyslexia on tasks assessing, for example, phonemic awareness (Bruck, 1992). However, there is evidence that individuals with developmental dyslexia show less categorical perception of phonemic contrasts (Serniclaes, Sprenger-Charolles, Carré, & Démonet, 2001; Serniclaes, Van Heghe, Mousty, Carré, & Sprenger-Charolles, 2004), suggesting over-specified phonological representations. The debate about the nature of the phonological deficits between underspecified vs. over-specified phonological representations is difficult to resolve by using traditional standardized measures, such as phonological awareness and rapid automatized naming, because they reveal the existence of phonological deficits without pinpointing the underlying sources of such impairment. Therefore, we used a more fine-grained eye-tracking method to investigate subtle individual differences in the sensitivity to subphonemic information. In addition, the natural variation of phonological skills due to environmental factors in the typically developing population has been understudied. This is particularly true for adults, since most reading research focuses on children. We recruited participants from a community sample of adults who had never been diagnosed with dyslexia or other learning disability, but nonetheless varied widely in reading and other abilities. This sample is more representative of the range of socio-economic status (SES) and

cognitive abilities in the general population than college student samples typical of much psycholinguistic research.

Overview: Phonological Skills and Reading Abilities

In the literature about the cognitive processes involved in reading, it has been widely accepted that phonological processing mediates the mapping between print and meaning (Harm & Seidenberg, 2004; Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979; Ziegler & Goswami, 2005). For example, in typical reading acquisition, pre-reading children's phonological processing abilities can predict their later reading and spelling abilities (Sprenger-Charolles, Siegel, Béchennec, & Serniclaes, 2003). More importantly, phonological processing abilities continue to contribute to reading and spelling abilities, in novice and expert readers alike, even when orthographic processing abilities are taken into consideration (Sprenger-Charolles et al., 2003). It has also been demonstrated with functional neuroimaging data that the amount of overlap between neural substrates of speech and print increases with reading skills (Shankweiler et al., 2008), further supporting the idea that phonological ability is an important locus where good and poor readers differ. In addition to typical reading acquisition, phonological deficits have been implicated in developmental reading disorders, such as dyslexia (Liberman, 1973; Ramus, 2003). For example, individuals with dyslexia tend to have difficulties in a range of phonological processes, including phonological awareness (Bruck, 1992), phonological short-term memory (McDougall, Hulme, Ellis, & Monk, 1994), and pseudo-word naming (Sprenger-Charolles, Colé, Lacert, & Serniclaes, 2000).

In addition, consistent with the high reliability of the link between phonological skills and reading abilities found both in studies targeted at typical and at atypical reading acquisition, it

has been suggested that phonological processing skills modulate the outcome of reading acquisition in a continuous manner across typical and atypical readers (Snowling, Gallagher, & Frith, 2003; Snowling & Hayiou-Thomas, 2006). Thus, knowing how exactly reading ability varies as a function of phonological processing could have a strong and wide impact on various theoretical and practical issues, such as verifying theories, modifying computational models, and improving educational and interventional strategies. Currently, variation in the nature of phonological processing has mostly been studied with individuals with dyslexia, while it is understudied in typical reading development. Yet, findings based on individuals with dyslexia could shed some light on what may also apply to the typical population. Therefore, we now turn to two of the most commonly investigated phonological theories of dyslexia to lay the foundations for our further discussion of typical reading acquisition.

Theories of Phonological Deficits in Dyslexia

Although a central and causal role of phonological deficits in dyslexia has been consistently supported by different research, there are still disagreements regarding the specific nature of phonological issues that lead to reading deficits (Ramus, Marshall, Rosen, & Van Der Lely, 2013). Two of the most debated phonological deficits theories are underspecified vs. over-specified phonological representations (Noordenbos, 2013).

Theories that appeal to underspecified phonological representations in dyslexia reason that the phonological deficits originate from impaired sensitivity to the acoustic changes in speech stimuli (Tallal, 1980; Tallal et al., 1998). Underspecified phonological representations indicate less distinctness of a phonological representations from its neighbors (Elbro, 1998), such that fuzzy boundaries between phonemes are the driving force of the difficulty in the grapheme-

phoneme mapping. Furthermore, it has been found that measures for distinctness of phonological representations can predict pre-reading children's future reading abilities. For example, Elbro, Borstrøm, and Petersen (1998) found that kindergarteners who produced less distinct pronunciations were significantly more likely to develop dyslexia in the future, even when their non-verbal IQ, articulatory fluency, and lexical access were taken into account.

On the other hand, perhaps less intuitively, the observed phonological deficits may instead stem from *over*-specified phonological representations. This line of reasoning argues that those with dyslexia maintain higher sensitivity to subphonemic details of native and non-native phonetic contrasts across the board, suggesting their phonological representations have not undergone the same degree of reorganization that seems to accompany reading acquisition in typical readers (Serniclaes, 2006). For example, studies have shown less categorical perception in dyslexia, demonstrated by reduced between-category discrimination and enhanced within-category discrimination, indicating that dyslexics are more sensitive to variants of the same phoneme (Serniclaes et al., 2001) and to allophonic variants (Serniclaes et al., 2004). While understanding speech with an allophonic "grain size" (what Serniclaes calls "allophonic perception") may not cause too much problem, the mismatch between spoken categories and graphemes may cause important problems in reading acquisition (Serniclaes, 2006). Given the profound perceptual changes required by reconstructing the mapping between sounds and categories from allophonic perception to phonemic perception, natural variation in the extent of this process during development may be not only due to genetic factors but also environmental ones (Serniclaes, 2006). That is, reading development may be impeded not just by biological predispositions, but by environment and experience; some children may not reach their potential

as readers due to specifically phonological factors that arise as a consequence of limited experience or poor instruction.

Gaps in the literature

Despite a long history of investigating phonological processes in developmental dyslexia, there is still no clear consensus as to the nature of these phonological deficits. That is, the debate between the accounts of underspecified and over-specified phonological representations in dyslexia has not been resolved. Most assessments used to probe the phonological processing in poor readers have been standardized tests, such as phonological awareness and rapid automatized naming, that point to the existence of phonological deficits without revealing the underlying sources of such impairment. Thus, to tell these two theoretical accounts apart, measures that are more fine-grained than commonly used standardized tests are needed to observe the individual differences in subphonemic processing. The Visual World Paradigm (VWP), an eye-tracking method, has proved fruitful in measuring the fine-grained nature of online speech processing at various linguistic levels, including discourse (Altmann & Kamide, 2009; Engelhardt, Bailey, & Ferreira, 2006), syntactic (Chambers, Tanenhaus, & Magnuson, 2004; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), semantic (Huettig & Altmann, 2005; Kaiser, Runner, Sussman, & Tanenhaus, 2009), lexical (Magnuson & Nusbaum, 2007), phonemic (Allopenna, Magnuson, & Tanenhaus, 1998; Magnuson, Tanenhaus, Aslin, & Dahan, 2003), and subphonemic levels (Dahan, Magnuson, Tanenhaus, & Hogan, 2001; McMurray, Aslin, Tanenhaus, Spivey, & Subik, 2008). Although, compared to reading abilities, general speech perception and comprehension does not seem to be severely affected by apparent phonological differences in dyslexia, it could be compensated by many of the other cues in speech signal, such

as context and prosody (McQueen & Cutler, 2001). Therefore, subtle differences at subphonemic and phonemic levels may still be observable in spoken word recognition with the VWP, when no additional higher-level cues are provided. The differences observed in spoken word recognition processing may further inform us about how differences at the word level, cascaded from the subphonemic and phonemic levels, can influence individuals' reading abilities.

Another concern is that the vast majority of the literature divides subjects into dichotomous groups, that is, typical vs. dyslexic readers. However, in reality, language abilities are continuously distributed in the population, not dichotomously. Indeed, studies that compared analyses between dichotomous and continuous approaches often find better statistical model fitness when treating language ability as a continuous predictor (e.g., language impairment; McMurray, Munson, & Tomblin, 2014). Thus, in order to understand the subtle individual differences in the relationship between phonological and reading skills, a continuum method of analysis needs to be applied.

Furthermore, the relationship between phonological skills and reading abilities has been mostly studied in the context of early reading acquisition. This type of study often stresses children's pre-school phonological processing abilities which in turn influencing their reading acquisition trajectory. However, little has been studied regarding environmental factors (e.g., failure of instruction at the home and/or school system) that could be compounded with genetic factors. In addition, limited attention has been drawn to the population at the endpoint of development, i.e., adults, who may have developed poor phonological and reading skills due to a combination of genetic and environmental factors, but have been overlooked by the educational and clinical systems. Therefore, there is an urgent need to understand and characterize reading-related abilities in low-literacy adults in order to address this neglected public health issue.

The Current Study

To address the gaps in the literature mentioned above, the present study investigated individuals' sensitivity to subphonemic information using an eye-tracking task in young adults with varying reading abilities. We used the eye-tracking paradigm in the study of Dahan et al. (2001). Dahan et al. (2001) extended the basic VWP for spoken word recognition (Allopenna et al., 1998) to subcategorical, that is, subphonemic, details in speech. In the basic paradigm, as participants listen to the spoken stimuli, participants' eye movements to pictures are assumed to reflect the real-time activation of the pictures' names. Dahan et al. (2001) were revisiting previous work by Marslen-Wilson and Warren (1994), who found a theoretically surprising pattern of results when participants heard items with misleading coarticulatory cues. Dahan et al. (2001) created spoken tokens with misleading coarticulatory information (or subcategorical mismatches) by splicing spoken words. For example, they took a target word (W1; e.g., 'cat') and spliced its final consonant onto the initial portion (up to the end of the vowel) of another token of W1, of a real word (W2; e.g., 'cab'), or of a non-word (N3; e.g., 'cag'). Thus, they had three forms of each target word: an identity-spliced token with no misleading coarticulation (W1W1; ‘ca[t]t’), a cross-spliced token with misleading coarticulation that favored a lexical alternative (W2W1; ‘ca[b]/t’), and a cross-spliced token with misleading coarticulation that did not favor a lexical item (N3W1; ‘ca[g]/t’).

A model like TRACE (McClelland & Elman, 1986) would predict that W2W1 should be harder to process than N3W1, because the initial portion of W2W1 matches a word (W2), which should be strongly activated and compete with W1, while the initial portion of N3W1 (N3) would not selectively activate a specific word. Counter to this prediction, Marslen-Wilson and Warren (1994) used a lexical decision task and found that W2W1 and N3W1 both took longer to

recognize than W1W1, but W2W1 was recognized just as quickly as N3W1. Dahan et al. (2001) asked whether the lexical decision task might not be sufficiently sensitive to detect differences. Using the VWP, they found that, in typical college students, the average looks to the target picture were the highest when there was no mismatch in the coarticulation (W1W1), lower when the subcategorical mismatch did not favor any lexical item (N3W1), and lowest when the subcategorical mismatch favored a lexical alternative (W2W1), even when the participants recognized all variants of the tokens as the name of the target picture (W1). Dahan et al. (2001) suggested that the difference between W1W1 and N3W1 reflects the phonological component of the subcategorical mismatch effect, whereas the difference between N3W1 and W2W1 reflects a lexical competition effect. That is, both N3W1 and W2W1 differ from W1 phonologically, whereas W2W1 adds the influence of a specific lexical competitor. Thus, this paradigm appears to index phonological processing at a very fine grain, and may provide a sensitive measure of phonological abilities in adult poor readers.

We utilized this eye-tracking paradigm to address the following questions. Does sensitivity to subphonemic information differ as a function of reading abilities? If so, does sensitivity to subphonemic information decrease or increase as the reading abilities decrease, indicating underspecified or over-specified phonological representations, respectively?

Furthermore, considering the low variation in reading abilities among the typical college population, the current study recruited college-aged young adults with a wider range of reading abilities from a community sample in order to employ the individual differences analysis. Moreover, the current sample also includes a lower range of socioeconomic status (SES) compared to typical college samples, suggesting that the individual differences observed in the present study may be heavily influenced by environmental factors, given that low SES

individuals are at a disadvantage in terms of language experience from infancy onward (Aikens & Barbarin, 2008; Coley, 2002).

Predictions

First, to replicate the well-established link between phonological skills and reading abilities, we predicted that individuals' performance on standardized phonological processing tasks would be highly correlated with their performance on reading comprehension tasks. This approach not only addressed individual differences, but was also intended to reveal the potential influence of environmental factors on phonological and reading abilities by examining low SES adult readers who had never been diagnosed with dyslexia (though it is possible that some might have met criteria had they been carefully assessed as school children).

Second, given that the eye-tracking task tapped into an individual's phonemic and subphonemic processing, we predicted that the individuals' phonological skills would be highly correlated with the subphonemic mismatch effect observed in their eye-tracking responses. This individual differences analysis was used to examine whether and how the subtle differences of phonological processing in readers could indeed be manifested in the process of spoken word recognition.

Third, assuming there would be a strong correlation between individuals' phonological skills and their subcategorical mismatch effect, we could predict two completely opposite directions of correlations based on the two types of phonological deficits that have been posited in the literature, that is, underspecified vs. over-specified phonological representations. If individuals' phonological deficits stem from having underspecified phonological representations, one would predict that poorer readers' subcategorical mismatch effect should be smaller than that

of better readers. Conversely, if individuals' phonological deficits originate from over-specified phonological representations, poorer readers would have greater subcategorical mismatch effects than better readers.

Finally, we predicted that individuals' phonological processing skills would also be correlated with the size of the lexical effect (i.e., the difference between N3W1 and W2W1) observed in the eye-tracking data. Though not central to the main interests of the current study, it was assumed that readers' quality of lexical representations varies with their reading ability, which is tightly related to their phonological processing skills. Therefore, we predicted that better readers (or better phonological skills performers) would have greater lexical effect due to stronger competition from the alternatives in the lexicon.

Methods

The data in the current study were collected as part of the larger study conducted by Braze, Shankweiler and colleagues that investigated individual differences of language and reading learning in young adults (e.g., Braze et al., 2016; Braze, Tabor, Shankweiler, & Mencl, 2007; Kukona et al., 2016), and a preliminary report with a subset of the current sample (N = 32) was reported by Magnuson et al. (2011).[1]

[1] Note that I joined the team after data had been collected by Braze and colleagues at Haskins Laboratories. My role has been to conduct the complete data analysis and theoretical interpretation, in consultation with Braze and Magnuson.

Participants

The participants were 67 college-aged native English speakers, whose age ranged from 16 to 25 years (*M* = 20.93; *SD* = 2.13). The participants had 11.81 years of education on average (*SD* = 1.51) and were recruited from community colleges and GED programs in the New Haven area. The recruitment included, by design, individuals with a wide range of backgrounds as well as cognitive and reading abilities. Although the current community sample has more individuals at the low end of the reading ability continuum, they did not report having diagnosed reading or learning disabilities. The participants gave informed consent and received financial compensation for their participation. All protocols were approved by the Yale University Human Investigation Committee. Four participants were excluded from analysis due to failing to complete significant portions of the tasks, resulting in 63 participants for further analysis.

Materials

Subcategorical Mismatch Task. The auditory materials were those originally used by Dahan, Magnuson, Tanenhaus, and Hogan (2001) and consisted of 15 triplets of one target word (W1), one competitor (W2) and one non-word (N3) that shared the same onset, such as *cat*, *cab* and *cag*, respectively (for the full set of the 15 triplets, see Appendix 1). The materials are further manipulated within each triplet by splicing the final stop consonant of W1 onto the initial portion (up to the end of the vowel) of another token of W1 (W1W1; e.g., 'ca[t]t'), of a token of W2 (W2W1; e.g., 'ca[b]t'), or of a token of N3 (N3W1; e.g., 'ca[g]t'). The visual materials were similar to those used in Experiment 2 in Dahan et al. (2001), except that their line drawings were replaced with photographs. See Appendix 2 for the full list of the visual materials.

Linguistic and Cognitive Abilities Assessment Battery. This study was part of a larger study of language abilities in community samples of adolescents and young adults. More than 30 individual difference measures were collected for the overall study. For the purposes of our analyses, we have selected a subset of 27 measures of various linguistic abilities, cognitive abilities, and demographic indicators based on the selection in Kukona et al. (2016), with the following exceptions. First, we did not include the anti-saccade task (Muñoz, Everling, Munoz, & Everling, 2004) nor the working memory assessment (Daneman & Carpenter, 1980) because these tasks were not administered on either a large portion or all of the participants. Second, we included socioeconomic status (Hollingshead, 1975) given our particular interest in environmental factors' influence on phonological development and reading acquisition. The majority of these measures were standardized assessments widely used in clinical and educational settings, as well as in the psycholinguistic literature. Each of the assessments in the battery was categorized as one of the following: reading comprehension, listening comprehension, vocabulary, decoding, reading fluency, rapid automatized naming, phonological skills, print experience, general cognitive abilities, and demographics.

Reading comprehension assessments included the Gates-MacGinitie Reading Tests, Fourth Edition (GM; MacGinitie, MacGinitie, Maria, & Dreyer, 2000), odd-numbered items of the Reading Comprehension subset in the Peabody Individual Achievement Test, Revised (PIAT; Braze et al., 2007; Markwardt, 1989), the Fast Reading subtest of the Stanford Diagnostic Reading Test, Fourth Edition (SDRT; Karlson & Gardner, 1995), and the Passage Comprehension subtest of the Woodcock-Johnson-III Tests of Achievement (WJ; Woodcock, McGrew, & Mather, 2001). Listening comprehension assessments included tape-recorded, even-numbered items of the Reading Comprehension subset in PIAT to assess both reading and

listening comprehension with well-matched tasks (Braze, Tabor, Shankweiler, & Mencl, 2007) and the Oral Comprehension subtest of the WJ. Vocabulary assessments comprised the Peabody Picture Vocabulary Test, Revised (PPVT; Dunn & Dunn, 1997) and the Vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999).

Furthermore, word decoding skills were assessed with the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (TOWRE; Torgeson, Wagner, & Rashotte, 1999) and the Letter-Word Identification subtest of the WJ, and non-word decoding was assessed with the Phonemic Decoding Efficiency subtest of TOWRE and the Word Attack subtest of the WJ. Reading fluency was evaluated by three passages from Gray Oral Reading Test, Fourth Edition (GORT; Wiederholt & Bryant, 2001) and the Reading Fluency subtest of the WJ. Rapid automatized naming (RAN) was assessed via the three Rapid Naming subsets (i.e., Colors, Digits, and Letters) of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999). Assessments of phonological skills included the phonological awareness (i.e., elision and blending) and phonological memory (i.e., digits and non-word repetitions) tests of CTOPP. Print experience was assessed by recognition of author and magazine names (Stanovich & Cunningham, 1992). General cognitive abilities assessments included visuospatial memory with Corsi Blocks (Corkin, 1974), matrix reasoning (WASI), and full-scale IQ (WASI; the two-subset form consisting of Vocabulary and Matrix Reasoning). Finally, demographics information included age, years of education, and socioeconomic status.

Procedure

The experimental eye-tracking task and the assessments were administered individually for each participant on two separate days, with about 3.5 hours per session. Breaks were

provided when necessary. Standard administration procedures and instructions were used for most published assessments, except that the Reading Comprehension subtest in PIAT was used for both reading and listening comprehension as described above (following the procedure described by Braze et al., 2007). The procedure for the visual world task was identical to that of Experiment 2 in Dahan et al. (2001), except that the materials were presented on a desktop computer and the eye movements were tracked using an SR-Research Eyelink II head-mounted eye tracker, sampling at 250 Hz; details follow. Participants were randomly assigned to one of the 3 lists, varying in which 5 target words were assigned to each of the three conditions, i.e., W1W1 (consistent coarticulation), W2W1 (misleading cohort coarticulation), and N3W1 (misleading non-word coarticulation). There were 30 trials in total, with 15 experimental trials (i.e., 5 for each condition) and 15 filler trials. On each trial, a fixation cross appeared on the screen along with four simple geometrical shapes. The trial began when the participant clicked the cross, and pictures of four objects appeared, including one target (e.g., a cat), one competitor (e.g., a cab), and two unrelated distractors (e.g., a vase and a tree), along with four geometric shapes (see Figure 1 for an example). Participants were instructed to use a computer mouse to follow spoken instructions presented via speakers, such as "Point to the vase. Now the cat. Now click on it and put it below the circle." Eye movements were recorded throughout each trial starting from the click on the fixation cross till the completion of the trial.

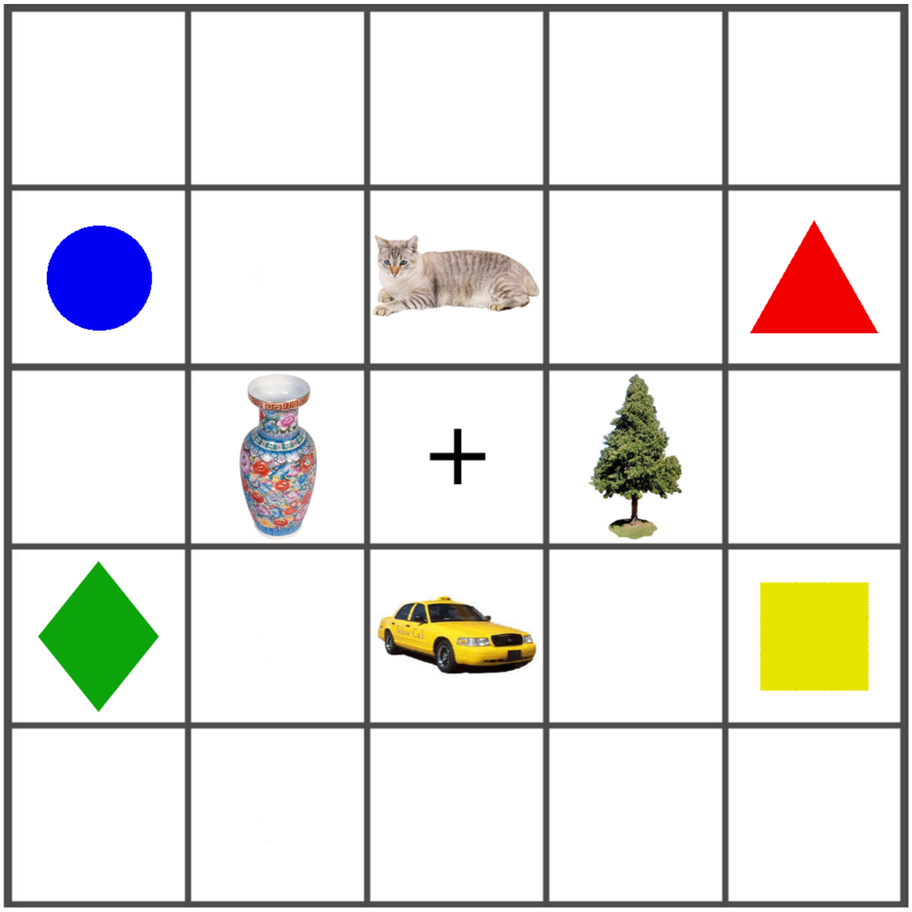

Figure 1. An example of the visual display during the eye-tracking experiment. The locations of the experimental pictures (i.e., target, competitor, and unrelated items) were randomized across trials and participants among the following cells: above, below, on the left of, and on the right of the center cross. The locations of the four geometric shapes were fixed across trials and participants in the positions shown in the figure.

Results

All statistical analyses were conducted using packages in the R statistical environment version 3.3.1 (R Core Team, 2016).

Individual Difference Measures

Due to missing data in a few measures (three missing values among the 63 participants across three measures, i.e., the two Reading Fluency measures and the SDRT Reading Comprehension measure), multiple imputation was applied to the dataset to replace the missing values with the imputed values using the *mice* package (van Buuren & Groothuis-Oudshoorn, 2011) before further analysis. For most measures, higher scores indicated better performance. Exceptions are the three sub-tests of CTOPP Rapid Automatized Naming (Colors, Digits, and Letters). The direction of the scoring scales was regularized across individual difference measures before further analysis. The raw scores of the exceptional measures (i.e., higher scores indicate worse performance) were transformed by subtracting participants' scores from the maximum observed score of the corresponding measure.

We observed skewness in most of the raw-score distributions based on the Q-Q plot of each assessment. Thus, Box-Cox transformations were applied to all assessment scores to normalize the distributions before further analysis (Box & Cox, 1964). Optimal lambda values were identified using the *boxcox* function from the *MASS* package (Venables & Ripley, 2002), and the transformations were carried out using the *bcpower* function from the *car* package (Fox & Weisberg, 2011). Box-Cox transformed scores were further standardized to account for their highly heterogeneous variances across variables. Outliers were identified by visually inspecting the Q-Q plot of each transformed variable and the output of *influencePlot* function from the car

package (Fox & Weisberg, 2011), which takes into consideration Studentized residuals, hat values, and Cook's distances. After removing two subjects due to their extreme scores on the TOWRE Word Naming task, Box-Cox transformation and standardization were applied to the raw scores for the remaining subjects. Visual inspection of the distributions suggested no more overly influential outliers, resulting in the final analysis set of data from 61 participants. The descriptive statistics of untransformed scores are listed in Table 1, excluding outliers and imputed values. Wide ranges of assessment scores across the board indicated high heterogeneity in the current sample, suitable for individual differences analysis. Simple correlations among the individual difference measures, Box-Cox transformed and standardized, are shown in Table 2.

Composite Scores

Individual difference measures tapped into several key reading-related skills: *reading comprehension* (measures 1-4 in Table 1 and Table 2), *listening comprehension and vocabulary* (5-8), *decoding and reading fluency* (9-14), *rapid automatized naming* (15-17), *phonological skills* (18-19), and *print experience* (20-21). These key skills were categorized based on previous published work that used similar community samples and individual difference measures as the current study, for being part of the same larger study, such as Braze et al. (2016) and Kukona et al. (2016). Composite scores were generated by averaging and then standardizing the transformed measures within each category. Listed in Table 3 are the correlations among the composites and additional simple measures of general cognitive abilities, i.e., matrix reasoning (measure 22 in Table 1 and Table 2), visuospatial memory (23) and WASI full-scale IQ (24). Since our main interest was individuals' phonological representations as a function of phonological skills, and individuals' phonological skills composite scores were highly correlated

with other reading-related assessment scores, the phonological skills composite was used as a proxy for overall reading ability for further analyses.

Table 1

Descriptive statistics of the individual difference measures.

Measures	N	M	SD	Range	Max
Reading Comprehension					
1. GM	61	29.93	9.61	10 - 47	48
2. PIAT	61	24.95	6.76	12 - 41	41
Grade Equivalent		5.86	2.61	2.5 - 13	-
3. SDRT	60	14.28	6.44	4 - 30	30
4. WJ	61	32.82	4.16	22 - 43	47
Grade Equivalent		7.55	4.47	2.4 - 19	-
Listening Comprehension					
5. PIAT	61	27.72	7.72	9 - 41	41
Grade Equivalent		7.06	2.93	2.1 - 13	-
6. WJ	61	23.95	3.73	17 - 32	34
Grade Equivalent		9.88	4.36	3.5 - 19	-
Vocabulary					
7. PPVT	61	158.97	18.85	116 - 197	204
8. WASI	61	45.03	12.11	17 - 78	66
Decoding					
9. TOWRE Words	61	87.79	8.95	68 - 104	104
10. WJ Words	61	63.31	6.29	49 - 75	76
Grade Equivalent		10.02	4.47	4 - 19	-
11. TOWRE Non-words	61	40.44	12.73	8 - 61	63
12. WJ Non-words	61	24.11	5.10	11 - 32	32
Grade Equivalent		8.20	4.90	2.3 - 19	-
Reading Fluency					
13. GORT	60	16.73	6.80	4 - 29	30
14. WJ	60	63.27	15.44	23 - 98	98
Grade Equivalent		9.74	3.84	2.6 - 19	-
Rapid Automatized Naming					
15. CTOPP Colors	61	39.43	7.51	27.2 - 60.9	-
16. CTOPP Digits	61	23.58	4.30	16.4 - 35.4	-
17. CTOPP Letters	61	25.07	4.33	18 - 37.4	-
Phonological Skills					
18. CTOPP PA	61	81.38	16.37	58 - 115	150
19. CTOPP PM	61	91.34	10.59	73 - 112	150

Print Experience					
20. Authors	61	3.18	3.69	-1 - 18	80
21. Magazines	61	5.25	4.52	-2 - 17	80
General Cognitive Abilities					
22. WASI Matrix	61	25.15	5.25	7 - 35	35
23. Corsi Blocks VM	61	4.79	1.09	2.2 - 7.2	9
24. WASI Full-Scale IQ	61	89.95	17.07	55 - 138	-
Demographics					
25. Age (Years)	61	20.96	2.20	16.88 - 24.8	-
26. Years of Education	61	11.74	1.48	8 - 16	-
27. SES	61	42.20	12.93	20 - 56	66

Note. N = sample size; M = mean; SD = standard deviation; Max = maximum possible score. GM = Gates-MacGinitie Reading Tests; PIAT = Peabody Individual Achievement Tests; SDRT = Stanford Diagnostic Reading Test; WJ = Woodcock-Johnson Tests of Achievement; PPVT = Peabody Picture Vocabulary Test; WASI = Wechsler Abbreviated Scales of Intelligence; TOWRE = Tests of Word Reading Efficiency; GORT = Gray Oral Reading Test; CTOPP = Comprehensive Test of Phonological Processing; PA = phonological awareness; PM = phonological memory; VM = visuospatial memory; and SES = Socioeconomic Status.

Table 2

Correlations among the individual difference measures (Box-Cox transformed and standardized).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.
Reading Comprehension																							
1. *GM*																							
2. *PIAT*	**.63**																						
3. SDRT	**.65**	**.54**																					
4. *WJ*	**.67**	**.56**	**.55**																				
Listening Comprehension																							
5. *PIAT*	**.69**	**.64**	**.62**	**.66**																			
6. WJ	**.71**	**.65**	**.60**	**.61**	**.75**																		
Vocabulary																							
7. PPVT	**.70**	**.62**	**.70**	**.69**	**.77**	**.68**																	
8. WASI	**.75**	**.65**	**.76**	**.69**	**.70**	**.62**	**.74**																
Decoding																							
9. TOWRE W	**.47**	**.46**	**.46**	**.55**	**.43**	**.35**	**.48**	**.54**															
10. WJ W	**.61**	**.56**	**.52**	**.65**	**.60**	**.62**	**.73**	**.61**	**.62**														
11. TOWRE NW	**.37**	**.40**	**.30**	**.47**	**.34**	**.28**	**.47**	**.46**	**.76**	**.68**													
12. WJ NW	**.43**	**.46**	**.29**	**.44**	**.36**	**.32**	**.51**	**.40**	**.59**	**.76**	**.83**												
Reading Fluency																							
13. GORT	**.51**	**.42**	**.56**	**.49**	**.32**	**.44**	**.58**	**.48**	**.70**	**.60**	**.55**	**.42**											
14. WJ	**.62**	**.46**	**.66**	**.56**	**.41**	**.46**	**.46**	**.65**	**.66**	**.41**	**.42**	**.27**	**.59**										
RAN																							
15. CTOPP Colors	**.31**	.07	**.27**	**.36**	**.25**	**.28**	.17	.20	**.37**	**.27**	.19	.21	**.28**	**.40**									
16. CTOPP Digits	-.09	-.05	.01	.01	.00	-.08	-.15	.01	**.57**	.07	**.33**	.13	.21	.22	**.30**								
17. CTOPP Letters	.10	.14	.19	.20	.12	-.03	.02	.23	**.62**	.11	**.42**	.21	**.41**	**.35**	**.33**	**.61**							
Phonological Skills																							
18. CTOPP PA	**.46**	**.44**	**.47**	**.63**	**.56**	**.58**	**.56**	**.56**	**.41**	**.68**	**.52**	**.52**	**.39**	**.35**	.17	-.09	.08						
19. CTOPP PM	**.52**	**.28**	**.32**	**.51**	**.37**	**.37**	**.32**	**.39**	**.50**	**.38**	**.41**	**.36**	**.39**	**.46**	.19	.18	**.29**	**.29**					
Print Experience																							
20. Authors	**.63**	**.53**	**.48**	**.56**	**.48**	**.42**	**.57**	**.58**	**.58**	**.50**	**.48**	**.46**	**.49**	**.65**	.19	-.02	.12	**.43**	**.49**				
21. Magazines	**.45**	**.48**	**.35**	**.44**	**.41**	**.38**	**.46**	**.58**	**.36**	**.48**	**.36**	**.30**	**.31**	**.39**	.03	.08	.04	**.31**	**.31**	**.51**			
General Cognitive																							
22. WASI Matrix	**.59**	**.55**	**.58**	**.57**	**.68**	**.66**	**.59**	**.54**	**.34**	**.51**	**.30**	**.35**	**.42**	**.40**	**.30**	-.05	.16	**.57**	**.39**	**.34**	.13		
23. Corsi	**.47**	**.38**	**.39**	**.40**	**.43**	**.44**	**.45**	**.49**	**.40**	**.46**	**.40**	**.37**	**.38**	**.43**	**.49**	.05	.18	**.49**	**.30**	**.35**	.11	**.55**	
24. Full-Scale IQ	**.67**	**.66**	**.72**	**.66**	**.72**	**.66**	**.72**	**.84**	**.53**	**.61**	**.45**	**.39**	**.53**	**.53**	**.25**	.13	**.28**	**.52**	**.34**	**.43**	**.47**	**.77**	**.48**

Note. N = 61. The three missing data points were replaced by imputed values using the *mice* package in *R* and the scales of the three CTOPP RAN subtests were inverted (by subtracting from their maximum observed scores) before conducting correlational analysis. Pearson correlation test: $|r| \geqslant .22$, $p < .1$; $|r| \geqslant .25$, $p < .05$; $|r| \geqslant .33$, $p < .01$; $|r| \geqslant .41$, $p < .001$. Bolded values indicate $|r| \geqslant .25$, $p < .05$.

Table 3

Correlation matrix among the composite scores.

		1.	2.	3.	4.	5.	6.	7.	8.
1.	Reading Comprehension								
2.	Listening Comprehension & Vocabulary	**.91**							
3.	Decoding & Reading Fluency	**.72**	**.66**						
4.	Rapid Automatized Naming	.20	.12	**.47**					
5.	Phonological Skills	**.67**	**.65**	**.69**	.22				
6.	Print Experience	**.67**	**.63**	**.63**	.11	**.55**			
7.	Matrix Reasoning	**.68**	**.69**	**.47**	.17	**.60**	**.27**		
8.	Visuospatial Memory	**.49**	**.51**	**.50**	**.31**	**.49**	**.26**	**.55**	
9.	Full-Scale IQ	**.81**	**.83**	**.63**	**.28**	**.54**	**.51**	**.77**	**.48**

Note. N = 61. Composite scores were calculated based on the Box-Cox transformed and standardized measures in Table 2 by averaging and standardizing the measures within each category, including reading comprehension (measures 1-4), listening comprehension and vocabulary (5-8), decoding and fluency (9-14), RAN (15-17), phonological skills (18-19), and print experience (20-21). Additional simple measures of general cognitive abilities, matrix reasoning (22), visuospatial memory (23), and full-scale IQ (24) were also included. Pearson correlation test: $|r| \geqslant .22, p < .1$; $|r| \geqslant .25, p < .05$; $|r| \geqslant .33, p < .01$; $|r| \geqslant .41, p < .001$. Bolded values indicate $|r| \geqslant .25, p < .05$.

Fixation Proportions of the Eye-tracking Task

Within trials, the fixation proportions to the object pictures were computed over time. The eye movements were sampled throughout every trial at the rate of 250 Hz, and were down-sampled to 20 Hz (50 ms time steps) for all further analyses. For each trial, at each time step beginning from target word onset, we determined fixation location as falling into one of these five categories: the target, the competitor, a distractor, the cross, or elsewhere. The fixation proportions of the five locations were then computed over trials by condition and by participant at each time step, excluding the filler trials and trials with incorrect mouse responses (see Figure 2A). Distractor proportions were divided by the number of distractors to result in the mean proportion of fixations to distractors. Note that stimulus driven eye-movements usually have a 200 ms delay relative to the stimulus time-course (Fischer, 1992; Viviani, 1990), and that the splice point is around 400 ms. Thus, the time period of interest was from 600 ms after word onset to 1200 ms, where the pattern of fixation proportions stabilized.

The overall target fixation proportions replicated the subcategorical mismatch effects seen in Dahan et al. (2001), where the participants looked to the target faster and to a greater extent when there was no mismatching coarticulatory information in the word (W1W1), with slower and lesser target fixation proportions when mismatching corarticulation corresponded to a non-word (N3W1), and even slower and lesser target fixation proportions when the mismatching coarticulation was consistent with a word (W2W1). Similarly, the overall competitor fixation proportions also replicated the findings in Dahan et al. (2001), where the rank order of the competitor fixation proportions was complementary to that of the target fixation proportions,

showing the highest competitor fixation proportions in W2W1, followed by N3W1, and the lowest competitor fixation proportions in W1W1.[2]

The fixation proportions to distractors did not differ across conditions while having a slightly higher fixation proportions around target word onset compared to the target and competitor fixations. This reflected the residual eye-movements to the distractors due to the instructions structure, where in each trial, the participant was asked to point to a distractor picture before pointing to the target picture. The overall fixation proportions to the cross and other regions on the screen also did not differ across conditions and did not change notably over time, providing a baseline and ensuring that participants were paying attention to the task and the objects on the screen.

---

[2] It is worth noting that, although target fixations and competitor fixations are usually complimentary, there are examples in the literature where authors appear to choose freely one or the other, allowing the possibility for choosing the one that yields a "stronger" result. In inspecting the data, we discovered an oddity in this data when we conducted a planned analysis by reading ability tertiles, with consistent patterns in competitors across tertiles but striking changes in target fixation patterns. Therefore, we conducted further analyses on target fixations, while acknowledging that our statistical results must therefore be considered provisional.

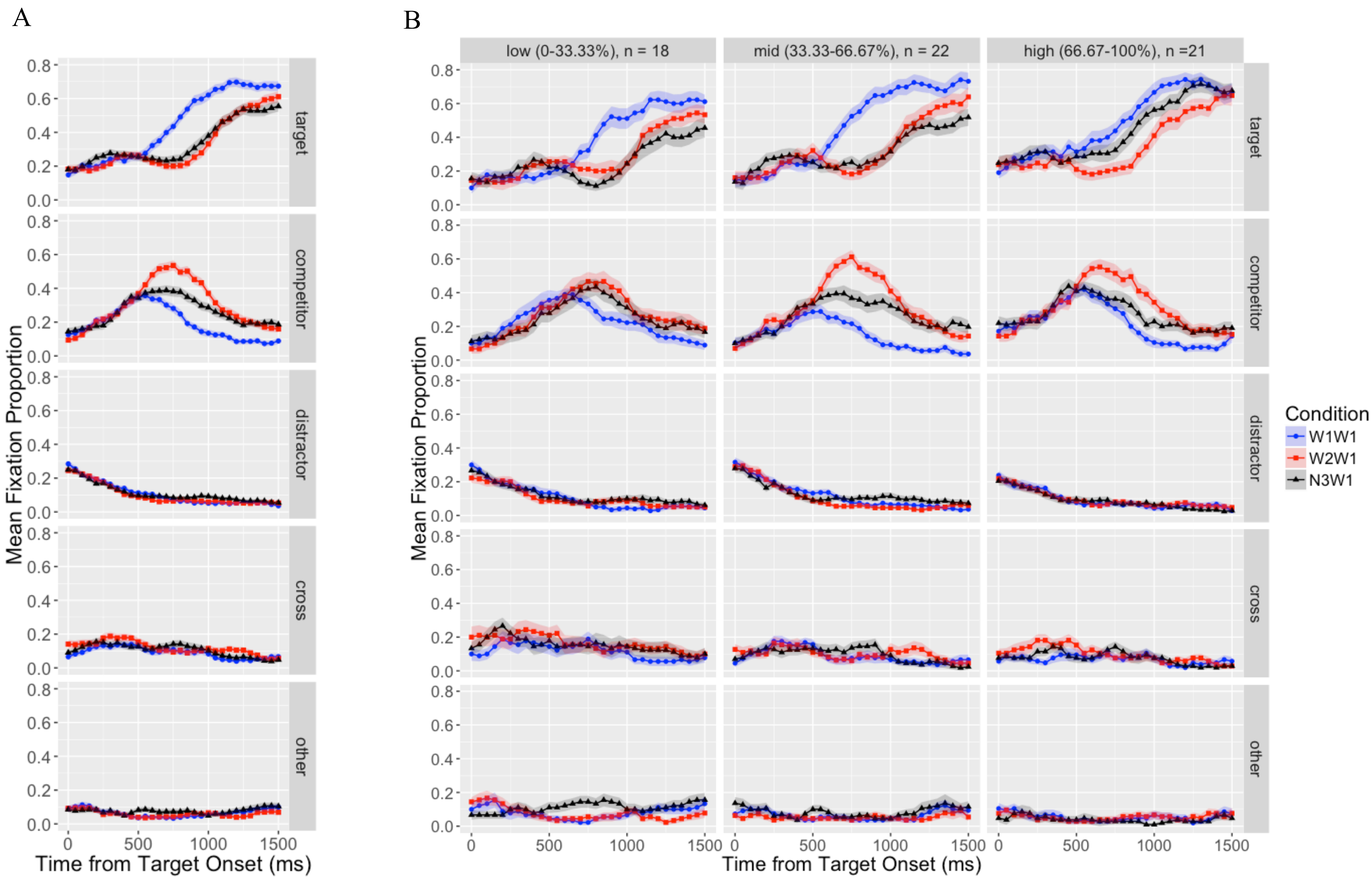

Figure 2. Mean fixation proportion by fixated object and by condition, (A) collapsed across all participants and (B) divided into tertiles of participants based on the phonological skills composite scores.

Growth Curve Analysis and Individual Differences

In order to characterize the individual differences in the eye-tracking data, we employed Growth Curve Analysis (GCA; Mirman, 2014) and extracted effect sizes (i.e., differences of target fixation proportions between conditions in a 600-1200 ms time window) for individual participants. We compared GCA effect sizes to assessment composite scores using correlational analyses. All GCA analyses were carried out using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) using a generalized linear mixed-effect model. Fixation proportion over time was modeled using orthogonal polynomial functions of time up to the third-order and fixed effects of conditions (i.e., W1W1, W2W1, N3W1) on all of the polynomial terms. The model also included random effects of participants on all temporal terms and that of participant-by-condition interaction on the intercept, linear and quadratic terms. Below is the model specification in *R*'s syntax:

```
meanFix ~ (ot1 + ot2 + ot3) * (CONDITION) +
(ot1 + ot2 + ot3 | SUBJECT) + (ot1 + ot2 | SUBJECT:CONDITION)
```

where `meanFix` is mean fixation proportion by condition, and `ot1`, `ot2`, and `ot3` are orthogonal time order 1 (linear), 2 (quadratic), and 3 (cubic), respectively.

For each participant, the participant-by-condition random effects estimates of the intercept were used to compute the effect sizes by subtracting the random effect estimate of N3W1 from that of W1W1 (i.e., the mismatch effect) and subtracting the random effect estimate of W2W1 from that of N3W1 (i.e., the lexical effect). The two effect sizes, i.e., the mismatch effect (W1W1-N3W1) and the lexical effect (N3W1-W2W1), were strongly and negatively correlated with each other $r(59) = .-55$, $p < .001$, indicating that participants whose mismatch effect was larger tended to have a smaller lexical effect, and vice versa. This suggests that,

individuals who have higher subphonemic sensitivity tend to have less lexical competition, possibly due to low-ability readers' poor lexical quality.

The correlation between the two effect sizes and the assessment composite scores were tested to further inspect the individual differences of language and other cognitive skills manifested in the eye-tracking data (shown in Table 4). Overall, the individual differences composite scores were negatively correlated with the mismatch effect (W1W1-N3W1) and positively correlated with the lexical effect (N3W1-W2W1). This suggests that participants whose language and cognitive abilities were worse tended to have higher subphonemic sensitivity and lower lexical competition, and vice versa. In addition, composite scores were generally more highly correlated with the lexical effect (N3W1-W2W1), suggesting that the individual differences were mainly reflected on the relative relationship between N3W1 and W2W1 conditions. The composite that showed the highest correlation to both the effect sizes was phonological skills, further supporting the choose of phonological skills composite as the individual differences indicator. Composites marginally to moderately correlated with the effects were reading comprehension, oral comprehension and vocabulary, decoding and reading fluency, rapid automatized naming, and print experience. The general cognitive skills (measures 7-9 in Table 4) were weakly correlated with the GCA effect sizes.

Table 4

Correlations between subcategorical mismatch effects and individuals' composite scores.

		W1W1-N3W1	N3W1-W2W1
	N3W1-W2W1	**-.55**	
1.	Reading Comprehension	-.12	.23
2.	Oral Comprehension & Vocabulary	-.21	**.25**
3.	Decoding & Reading Fluency	-.10	**.30**
4.	Rapid Automatized Naming	-.08	.20
5.	Phonological Skills	-.22	**.32**
6.	Print Experience	-.09	.20
7.	Matrix Reasoning	-.18	.11
8.	Visuospatial Memory	-.10	.19
9.	Full-Scale IQ	-.17	.24

Note. N = 61. Pearson correlation test: $|r| \geqslant .22$, $p < .1$, $|r| \geqslant .25$, $p < .05$; $|r| \geqslant .33$, $p < .01$; $|r| \geqslant .41$, $p < .001$. Bolded values indicate $|r| \geqslant .25$, $p < .05$.

Growth Curve Analysis with Phonological Skills as a Fixed Effect

To visually inspect the individual differences in the subcategorical mismatch effects observed in the eye movements, we divided the participants into tertiles based on their phonological skills composite scores, as shown in Figure 2B. The top one-third performers' target fixation proportions were very similar to the overall pattern qualitatively, in terms of the rank order of the conditions. Interestingly, as the phonological skills composite scores decreased, there was a trend for target fixation proportions to decrease in N3W1 but increase in W2W1, to the extent that individuals with poorer phonological skills had a reversal of rank order between W2W1 and N3W1. This reversal in the target fixations was completely unexpected, although poorer readers' heightened fixations in N3W1 to other regions on the screen (cf. bottom row of Figure 2B) could suggest that these participants may have noisier processing or that they may be more sensitive to the coarticulatory information and were searching for an alternative picture to match what they hear. We will discuss the reversal between W2W1 and N3W1 in more detail in the next section, The Effect of Place of Articulation in Coarticulation.

In order to quantify the effect of individual differences in phonological skills on the subcategorical mismatch effects, we applied GCA on the target fixation proportions again. The new GCA model was the same as the previous one, except that the phonological skills composite was now added as a fixed effect, together with its interactions with condition and time. Below is the model specification in *R*'s syntax:

```
meanFix ~ (ot1 + ot2 + ot3) * (CONDITION) * (PHONOLOGICAL) +
(ot1 + ot2 + ot3 | SUBJECT) + (ot1 + ot2 | SUBJECT:CONDITION)
```

The comparison between GCA models without and with the phonological skills composite as a fixed effect showed that adding the phonological skills composite into the model

significantly improved the model fit (Table 5), suggesting that individuals' phonological skills uniquely contributed to the variance in their eye-tracking patterns. We further examined parameter estimates for interactions involving phonological skills in order to assess individual differences in the timing and strength of lexical activation under conditions of cue ambiguity.

In order to examine all possible combinations of differences between conditions (i.e., W1W1 vs. N3W1, N3W1 vs. W2W1, and W1W1 vs.W2W1), we ran the same GCA model twice, one with N3W1 as the baseline and one with W1W1 as the baseline. This is because, the *lmer* function in the *lme4* package by default treats one of the conditions as the baseline, and the baseline condition in the model can be seen as the *control* or *reference* condition, which is used to estimate the differences between other conditions and the baseline (Mirman, 2014). Note that, changing the baseline would only provide different sets of parameter estimates, but not affect the overall model fit. The model fit is visualized in Figure 3A (collapsed across all participants) and Figure 3B (participants divided into tertiles), and parameter estimates of the model are listed in Table 6 (with N3W1 as the baseline) and Table 7 (with W1W1 as the baseline).

Table 5

Comparison between GCA models with vs. without the composite scores of phonological skills as a fixed effect.

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	p-value
without	29	-2906.9	-2739.4	1482.5	-2964.9			
with	41	-2914.1	-2677.3	1498.0	-2996.1	31.20	12	0.0018

Note. Adding phonological skills composite scores significantly improved the model fit. Df: degrees of feedom; AIC: Akaike information criterion; BIC: Bayesian information criterion; logLik: log-likelihood; Chisq: Chi-Square test value; Chi Df: Chi degrees of freedom.

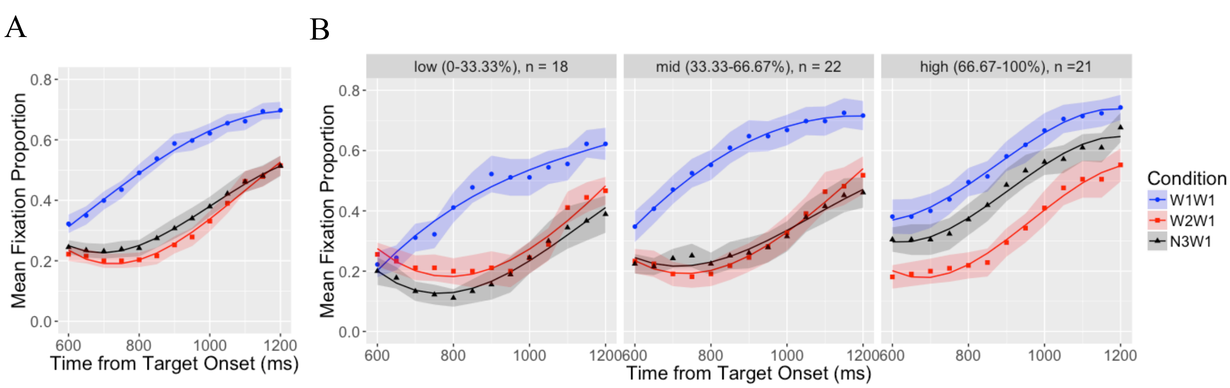

Figure 3. GCA model fit on target fixation proportions with participants as random effects and with conditions and phonological skills composite scores as fixed effects, (A) collapsed across participants and (B) divided into tertiles of participants based on individuals' composite scores of phonological skills. (cf. top row of Figures 1A and 1B, but note the difference in the time range)

N3W1 as the Baseline. The parameter estimates of the GCA model with N3W1 as the baseline are listed in Table 6. Overall, all four polynomial terms (intercept, linear, quadratic, and cubic) were statistically significant, indicating that the average target fixation proportion between W1W1 and W2W1 was greater than that of N3W1, and had a slope steeper and sharper points of inflections than that of N3W1. Effects of individual differences on the target fixation proportions were shown by the interactions between individual phonological skill composite scores and the polynomial terms. There were significant effects on the intercept term (*Estimate* = 0.098; *SE* = 0.021; $p < .001$) and the linear term (*Estimate* = 0.102; *SE* = 0.048; $p < .05$). This reflected that, as the individuals' phonological skill composite scores decreased, the average fixation proportion between W1W1 and W2W1 became greater relative to the fixation proportion of N3W1. While the results regarding the average between W1W1 and W2W1 are included and summarized here for the sake of completeness, we would like to draw attention to the following results, which are more central to the current study.

Among the parameters estimates of W1W1 on the polynomial terms, there was a significant effect of W1W1 on the intercept (*Estimate* = 0.206; *SE* = 0.027; $p < .001$) and on the quadratic term (*Estimate* = -0.171; *SE* = 0.044; $p < .001$). The intercept effect indicated that participants were more likely and faster to look to the target in the W1W1 than in the N3W1 condition. The quadratic effect reflected that the time course was less curved in W1W1 than in N3W1, where the target fixation proportion in N3W1 did not rise until 800 ms. On the other hand, there was no significant effect of W2W1 on any of the polynomial terms, suggesting that, on average, there was no significant difference in how much and how quickly the participants would look to the target picture between the W2W1 and the N3W1 conditions. (cf. Figure 3A)

Of the most interest was the interaction between the individual differences and the mismatch conditions over time (see Figure 3B). The individual differences in the phonological skills in W1W1 had marginally significant effects on the intercept (*Estimate* = -0.051; *SE* = 0.028; $p = .063$) and the quadratic term (*Estimate* = 0.079; *SE* = 0.044; $p = .072$), but no significant effects on the linear term (*Estimate* = -0.101; *SE* = 0.064; $p = .116$) nor the cubic term (*Estimate* = 0.019; *SE* = 0.017; $p = .272$). The individual differences in the phonological skills in W2W1 had a significant effect on the intercept (*Estimate* = -0.072; *SE* = 0.028; $p < .01$), but not the linear term (*Estimate* = -0.040; *SE* = 0.064; $p = .533$), the quadratic term (*Estimate* = 0.005; *SE* = 0.044; $p = .910$), or the cubic term (*Estimate* = 0.009; *SE* = .017; $p = .579$). Collectively, these significant interactions were consistent with the visual observations on the target fixation proportion curves, where when the phonological skills composite scores decreased, the differences between W1W1 and N3W1 increased (always positive values) while the differences between N3W1 and W2W1 decreased (from positive values to negative values). This suggests that poor readers show higher sensitivity to subphonemic information and lower lexical competition.

However, although using N3W1 as the baseline allowed us to observe both mismatch effect (W1W1-N3W1) and the lexical effect (N3W1-W2W1) in one model, there are a couple of important caveats. First, as the difference between N3W1 and W2W1 decreased, the so called “smaller” lexical effect became negative values, and there is no intuitive interpretation for a negative lexical effect based on our theoretical framework. Second, with N3W1 as the baseline, the difference between W1W1 and W2W1 could not be estimated, and thus it is not clear whether the relationship between W1W1 and W2W1 played a role in the correlation of the two effects. Therefore, we now turn to the same GCA model with W1W1 as the baseline.

Table 6

Parameter estimates of Growth Curve Analysis on individual differences effect of phonological skills composite on the subcategorical mismatch effects, using N3W1 as the baseline.

	Estimate	Std. Error	t value	p value	sig
(Intercept)	0.336	0.021	16.095	0.000	*
ot1	0.345	0.048	7.233	0.000	*
ot2	0.093	0.031	2.974	0.003	*
ot3	-0.047	0.017	-2.699	0.007	*
CONDW1W1	0.206	0.027	7.553	0.000	*
ot1:CONDW1W1	0.100	0.063	1.578	0.115	n.s.
ot2:CONDW1W1	-0.171	0.044	-3.919	0.000	*
ot3:CONDW1W1	0.032	0.017	1.898	0.058	n.s.
CONDW2W1	-0.031	0.027	-1.139	0.255	n.s.
ot1:CONDW2W1	0.027	0.063	0.420	0.674	n.s.
ot2:CONDW2W1	0.061	0.044	1.402	0.161	n.s.
ot3:CONDW2W1	0.004	0.017	0.242	0.808	n.s.
phono.composite	0.098	0.021	4.644	0.000	*
ot1:phono.composite	0.102	0.048	2.127	0.033	*
ot2:phono.composite	-0.058	0.031	-1.849	0.064	n.s.
ot3:phono.composite	-0.009	0.017	-0.528	0.597	n.s.
CONDW1W1:phono.composite	-0.051	0.028	-1.861	0.063	n.s.
ot1:CONDW1W1:phono.composite	-0.101	0.064	-1.572	0.116	n.s.
ot2:CONDW1W1:phono.composite	0.079	0.044	1.800	0.072	n.s.
ot3:CONDW1W1:phono.composite	-0.019	0.017	-1.098	0.272	n.s.
CONDW2W1:phono.composite	-0.072	0.028	-2.605	0.009	*
ot1:CONDW2W1:phono.composite	-0.040	0.064	-0.623	0.533	n.s.
ot2:CONDW2W1:phono.composite	0.005	0.044	0.114	0.910	n.s.
ot3:CONDW2W1:phono.composite	-0.009	0.017	-0.554	0.579	n.s.

Note. ot1 = linear term; ot2 = quadratic term; ot3 = cubic term. phono.composite = phonological skills composite scores. The normal approximation was used to compute parameter-specific *p*-values.

W1W1 as the Baseline. The parameter estimates of the GCA model with W1W1 as the baseline are listed in Table 6. When comparing the average fixation proportion between W2W1 and N3W1 to that of W1W1, the effects were statistically significant on the intercept, linear and quadratic terms. The intercept and linear effects reflected that the target fixation to W1W1 was greater and faster than that of W2W1 and N3W1. The quadratic effect reflected the visual observation that, while the target fixation was already rising at 600ms in W1W1, the average target fixation proportion of W2W1 and N3W1 did not start rising until 800 ms (see Figure 3A). This overall pattern indicates that participants were sensitive to the misleading coarticulatory information in W2W1 and N3W1, and thus looked to the target pictures less in these conditions. Among the parameters estimates of W2W1 on the polynomial terms, there was a significant effect of W2W1 on the intercept (*Estimate* = 0.237; *SE* = 0.027; $p < .001$) and on the quadratic term (*Estimate* = 0.233; *SE* = 0.044; $p < .001$), but not the linear (*Estimate* = -0.073; *SE* = 0.063; $p = .247$) nor the cubic terms (*Estimate* = -0.028; *SE* = 0.017; $p = .098$). This pattern was similar to the W1W1 effect on the polynomial terms when N3W1 was used as the baseline. The intercept effect indicated that participants were more likely and faster to look to the target in the W1W1 than in the W2W1 condition. The quadratic effect reflected that the time course was less curved in W1W1 than in W2W1, where the target fixation proportion in W2W1 did not rise until 800 ms. The N3W1 effect here is the same as the W1W1 effect with N3W1 as the baseline, except that the sign of direction is opposite for the parameter estimates. (cf. Figure 3A)

Effects of individual differences in phonological skills had a significant effect only on the intercept term (*Estimate* = 0.047; *SE* = 0.021; $p < .05$). This reflected that, as the individuals' phonological skill composite scores decreased, the average fixation proportion between W2W1 and N3W1 became lower relative to that of W1W1 (see Figure 3B), suggesting that poorer

readers had higher sensitivity to misleading coarticulatory information. The individual differences in the phonological skills in W2W1 had no significant effect on any of the polynomial terms, suggesting that the difference between W1W1 and W2W1 stayed fairly stable as a function of the phonological skills. Again, the N3W1 effect here is the same as the W1W1 effect with N3W1 as the baseline, showing marginally significant increase in difference between N3W1 and W1W1 as the phonological skills decreased. This suggests that the negative correlation between the mismatch effect (W1W1-N3W1) and the lexical effect (N3W1-W2W1), shown in Table 4, was driven mainly by participants' variation in N3W1, while the difference between W1W1 and W2W1 remained constant.

Table 7

Parameter estimates of Growth Curve Analysis on individual differences effect of phonological skills composite on the subcategorical mismatch effects, using W1W1 as the baseline.

	Estimate	**Std. Error**	**t value**	**p value**	**sig**
(Intercept)	0.542	0.021	25.970	0.000	*
ot1	0.445	0.048	9.332	0.000	*
ot2	-0.079	0.031	-2.520	0.012	*
ot3	-0.015	0.017	-0.855	0.392	n.s.
CONDW2W1	-0.237	0.027	-8.692	0.000	*
ot1:CONDW2W1	-0.073	0.063	-1.157	0.247	n.s.
ot2:CONDW2W1	0.233	0.044	5.321	0.000	*
ot3:CONDW2W1	-0.028	0.017	-1.655	0.098	n.s.
CONDN3W1	-0.206	0.027	-7.553	0.000	*
ot1:CONDN3W1	-0.100	0.063	-1.578	0.115	n.s.
ot2:CONDN3W1	0.171	0.044	3.919	0.000	*
ot3:CONDN3W1	-0.032	0.017	-1.898	0.058	n.s.
phono.composite	0.047	0.021	2.211	0.027	*
ot1:phono.composite	0.002	0.048	0.035	0.972	n.s.
ot2:phono.composite	0.021	0.031	0.674	0.500	n.s.
ot3:phono.composite	-0.028	0.017	-1.594	0.111	n.s.
CONDW2W1:phono.composite	-0.020	0.028	-0.744	0.457	n.s.
ot1:CONDW2W1:phono.composite	0.061	0.064	0.949	0.343	n.s.
ot2:CONDW2W1:phono.composite	-0.074	0.044	-1.687	0.092	n.s.
ot3:CONDW2W1:phono.composite	0.009	0.017	0.543	0.587	n.s.
CONDN3W1:phono.composite	0.051	0.028	1.861	0.063	n.s.
ot1:CONDN3W1:phono.composite	0.101	0.064	1.572	0.116	n.s.
ot2:CONDN3W1:phono.composite	-0.079	0.044	-1.800	0.072	n.s.
ot3:CONDN3W1:phono.composite	0.019	0.017	1.098	0.272	n.s.

Note. ot1 = linear term; ot2 = quadratic term; ot3 = cubic term. phono.composite = phonological skills composite scores. The normal approximation was used to compute parameter-specific *p*-values.

The Effect of Place of Articulation in Coarticulation

The GCA analysis demonstrated that the difference between W1W1 and N3W1 increased while the difference between N3W1 and W2W1 decreased when an individual had poorer phonological skills, indicating higher subphonemic sensitivity and smaller lexical competition effect in poorer readers. However, it is not intuitively clear as to why there would be a reversal of rank order between W2W1 and N3W1 in individuals with poorer phonological skills, because there is no theoretical or computational principle that would predict such pattern. W2W1 was expected to always attract fewer target fixations than N3W1, since it had been assumed that, compared to W1W1, N3W1 tokens had only phonological mismatch in the coarticulation, whereas W2W1 tokens had both mismatching phonological and lexical information embedded in the coarticulation.

Based on the GCA analysis (Table 7) and the visual inspection of the target fixation proportions when we divided participants into tertiles based upon the phonological skills composite scores (Figure 3B), it seems that it was mainly the target fixations in the N3W1 condition that were driving the differences we observed across participants along the phonological skills continuum. This led us to ask whether there might be some aspect of the stimuli that could explain the strange reversal of N3W1 and W2W1 rank orders among the lower-ability participants. Therefore, we conducted the following post-hoc exploratory analysis.

Recall that the original stimuli were designed such that W1-W2-N3 triplets were composed of syllables ending in a restricted set of consonants (in order to impose a degree of homogeneity and remove any phonetic bases for observed effects); final consonants were all stops with either labial (/b/ or /p/), alveolar (/d/ or /t/), or velar (/g/ or /k/) place of articulation. We asked whether it was possible that the consonants assigned to N3 and W2 in the triplets

might vary in their similarity to the consonants assigned to W1. Indeed, if we assume that labials and alveolars are more similar to each other than to velars due to greater similarity in (anterior) place of articulation, a possible confound becomes apparent[3]. We considered a triplet to have similar W1/N3 coarticulation when the final consonants of W1 and N3 were either labial or alveolar. We considered a triplet to have dissimilar W1/N3 coarticulation when one of the final consonants of W1 and N3 was velar, and the other was either labial or alveolar. Nine triplets fell into the W1/N3 similar category whereas six triplets fell into the W1/N3 dissimilar category (see Appendix 2 for more detail). If some subjects were more sensitive to phonetic similarity, might this modest difference be enough to induce the N3W1-W2W1 reversal observed in the lower tertiles?

Figure 4A shows the target fixation proportions based on W1/N3 coarticulation similarity (including all subjects). When the coarticulation between W1 and N3 was similar (cf. top panel of Figure 4A), the rank order of the three conditions was the same as the overall pattern, where W1W1 was greater than N3W1, followed by W2W1. However, when the coarticulation between W1 and N3 was dissimilar (cf. bottom panel of Figure 4A), the target fixations in N3W1 seemed to be suppressed to be at a similar level as W2W1, if not lower. The difference in N3W1 between the two categories suggests that the target fixations of cross-spliced conditions (N3W1 and W2W1) were heavily driven by the fine-grained subphonemic cues in the coarticulation, where mismatching items with coarticulation more similar to that expected from the target word elicited more target fixations. On the other hand, the pattern of the target fixations in W2W1 was

[3] Our classification of similarity in place of articulation still needs to be verified by acoustic analysis in the formant transition and/or by participants' subjective similarity rating. Note that, our classification is not consistent with some phoneme similarity metrics based on confusion matrices (e.g., Luce, 1987). However, it is very likely that the phoneme similarity reflected by confusion metrics of intact consonantal phonemes is heavily driven by the release of the consonants, whereas the coarticulation only contains the information of pre-release closure but not the release.

consistent with the lexical effect account. We assume both factors (lexical status and phonetic similarity) are in play in these results. More specifically, it is intuitive that, within W1/N3 similar items, the target fixations of W2W1 would be lower than N3W1 because there was an additive effect of lexical mismatch as well as dissimilar subcategorical mismatch. In contrast, within W1/N3 dissimilar items, the target fixations of W2W1 were slightly higher than that of N3W1, likely due to slight subcategorical mismatch plus lexical mismatch. Examining the target fixation pattern by item further supported the categorization based on W1/N3 coarticulation similarity (Figure 5), where 7 out of 9 items of the W1/N3 similar group showed the within-group average pattern, and 4 out of 6 items of the W1/N3 dissimilar group showed the within-group average pattern, suggesting that the average pattern for both groups of the items was valid and not driven by just a few items.

Furthermore, to visually inspect the individual differences in the W1/N3 coarticulation similarity effect, the target fixation proportions were further broken down into three participant groups in quantiles based on their phonological skills composite scores (Figure 4B). Participants with better phonological skills showed a pattern similar to the overall average pattern. Interestingly, participants with poorer phonological skills seemed to have a magnified phonological mismatch effect and a slightly smaller lexical effect (cf. top row of Figure 4B). In particular, the target fixations of the cross-spliced conditions were generally more suppressed in poorer phonological skills performers, indicating that they might be more sensitive to the subphonemic difference in the coarticulation than the better phonological skills performers. In addition, heightened W2W1 target fixations in W1/N3 dissimilar items among poorer phonological performers suggested that the mismatch in phonological information seemed to affect them more than the mismatch in lexical information (cf. bottom row of Figure 4B). We

will discuss implications of this apparently greater sensitivity to subphonemic detail among participants in the lower tertiles in the Discussion.

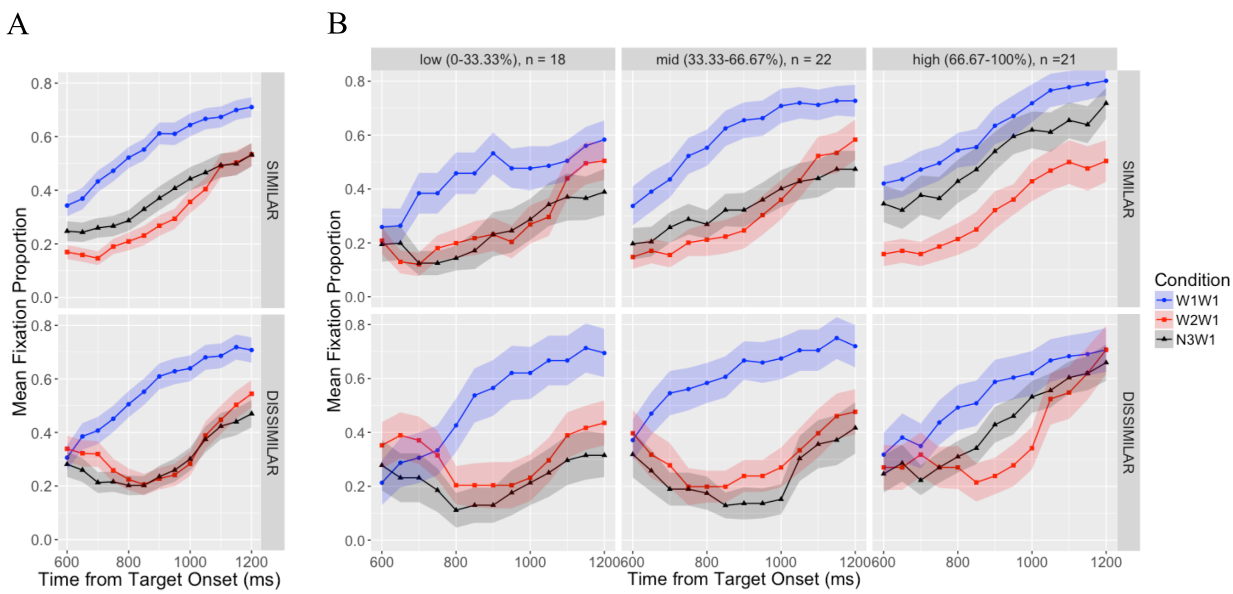

Figure 4. Target fixation proportions divided by place of articulation similarity between the coarticulation of W1W1 and of N3W1, (A) collapsed across all participants and (B) divided by into tertiles based on individuals' composite scores of phonological skills.

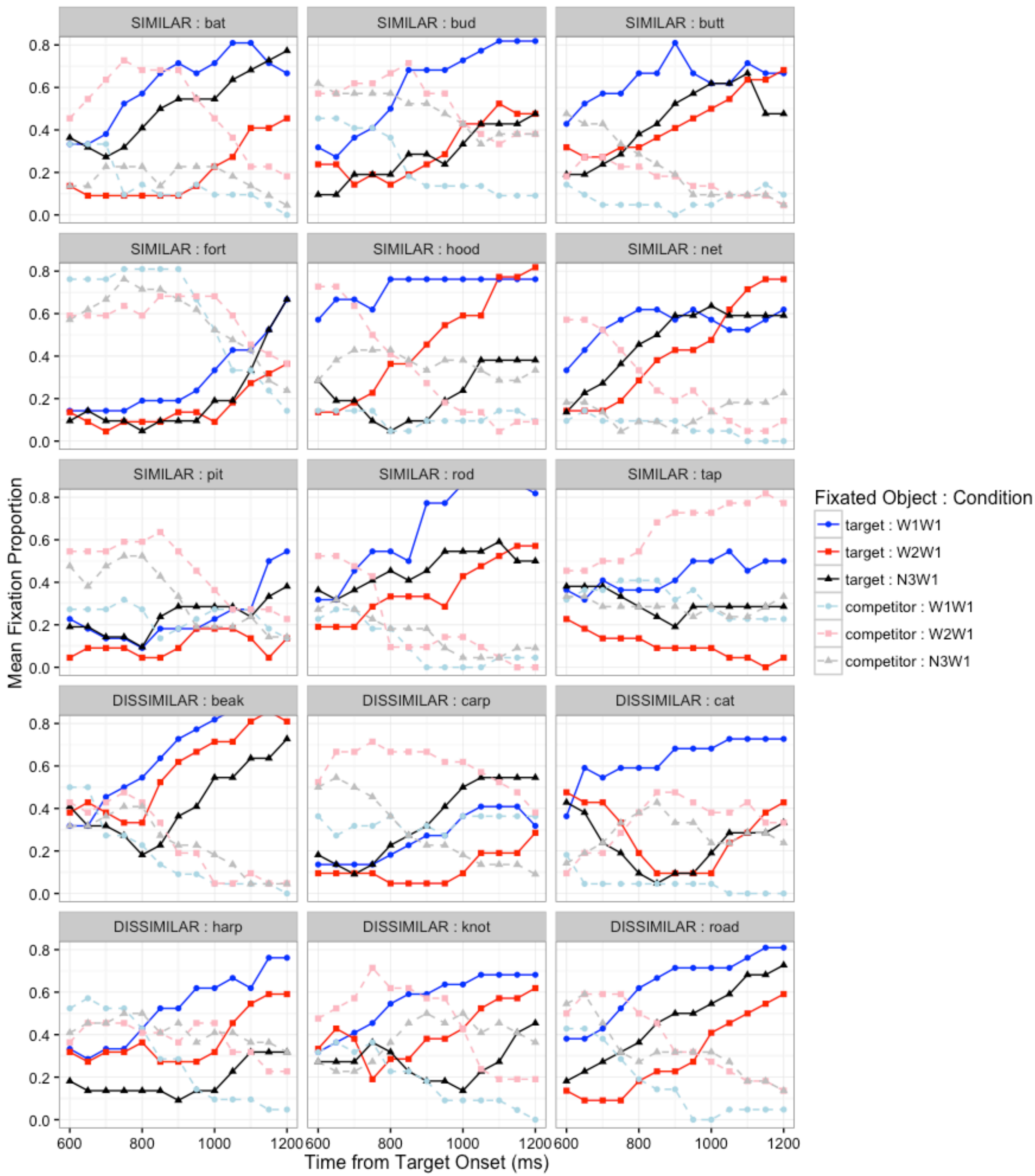

Figure 5. Target fixation proportions divided by place of articulation similarity between the coarticulation of W1W1 and of N3W1, separated by item.

Discussion

The present study investigated variation of sensitivity to subphonemic information as a function of reading abilities in young adults. Our analyses of standardized assessments replicated previous findings that individuals' reading performance varies with phonological skills. Results from our experimental task, the eye tracking version of the subcategorical mismatch paradigm (Dahan et al., 2001), add potentially crucial new information to individual differences in phonological and reading skills measured in standardized tests: specifically, lower-ability readers appear to exhibit higher sensitivity to subphonemic (phonetic) detail, consistent with the allophonic perception in dyslexia proposed by Serniclaes (2006).

The central tendency in our results replicated the findings of Dahan et al. (2001): participants' fixations to targets were slowed by mismatching coarticulation, with greater slowing when misleading coarticulation was consistent with a competitor word (W2W1 condition) than when it was consistent with a non-word (N3W1 condition). Higher phonetic sensitivity among poorer readers manifested most saliently in an unexpected reversal of N3W1 and W2W1; that is, poorer readers showed greater interference from coarticulation consistent with a non-word – a result that does not appear consistent with any extant theory or model of spoken word recognition. However, after taking into account the similarity of the place of articulation[4] embedded in misleading coarticulation in the vowel, the unexpected reversal effect could be explained by the higher phonological sensitivity in the poorer readers. It is possible that the smaller lexical competition effect observed in poorer readers also contributed to the reversal pattern. These findings point to important theoretical as well as translational implications.

[4] Presumably, the similarity of place of articulation would be reflected by the patterns in formant transitions, although we have not yet succeeded in uncovering the acoustic-phonetic basis for place perception in our materials.

To begin with, the insight from our study could facilitate refining current cognitive theories and models for reading development by taking into account both intrinsic and extrinsic factors. Currently, most of the cognitive models for reading acquisition have strived to understand the underlying neurobiological mechanisms causing reading disabilities, while environmental aspects are usually only vaguely mentioned or left out of the picture. For example, Harm and Seidenberg (1999) simulated the variation in phonological and reading acquisition by lesioning the connections in their connectionist model, representing the neurobiological alterations in individuals with dyslexia. In addition, in Harm and Seidenberg's (1999) model, there was no external parameter that was manipulated except the amount of time where the model was exposed to the input. In order to simulate a younger control group, the accumulated amount of exposure time was used to approximate the age of an individual, assuming that all individuals had the same consistent language input throughout their development. However, such an assumption may not hold true for a diverse and realistic population, as evidenced by previous literature indicating that the amount of spoken and written language input varies substantially based on individuals' socioeconomic status (Aikens & Barbarin, 2008; Coley, 2002). Thus, it is crucial to incorporate environmental factors, such as how frequently language input is present, into the model to estimate how the phonological representations would develop as a function of the richness of the language environment.

One way that Harm and Seidenberg (1999) evaluated the phonological acquisition of the models was to look at the degree of categorical perception the models achieved after phonological training. Thus, this model is ideal to simulate the specificity of phonological representations throughout language development, and, with slight modifications, it may account for the influences from environmental factors. For instance, a way to implement the variation in

language input among typically developing children at a similar age is to vary how often the stimuli are presented to the model while keeping the total training time constant during the training stage of phonological acquisition. Presumably, the model would achieve different degrees of categorical perception of phonology given the accumulated amount of input and the presentation rate, which simulates the phonological representations of individuals with various degrees of language input from the environment. Following the phonological acquisition stage, the training stage of reading acquisition of the model would result in different degrees of accuracy in phonology-orthography mapping, given the phonological representations from the previous training stage. The outcome of the training stage of reading acquisition would inform us how reading abilities vary with the phonological representations.

Furthermore, if our arguments can be supported by the results of the computational modeling suggested above, together with our current experimental findings, it highlights an important public health issue that may have been neglected. That is, individuals that are deprived from rich language environment can be detrimentally affected in a way that is similar to being genetically predisposed to neurological based language disorders, such as dyslexia. It is probable that the reading and phonological abilities of our low SES sample were affected by a mixture of genetic (dyslexia) and environmental (failures of instruction) factors. However, given the continuous profile of phonological skills and relatively higher proportion of lower-ability readers in our sample (the prevalence of dyslexia is 5-17.5%; Lagae, 2008), it is very likely that the environmental factors contributed strongly to this variation.

Although research has shown that low SES communities are under-resourced both in the home and the school system (Aikens & Barbarin, 2008; Muijs, Harris, Chapman, Stoll, & Russ, 2004; Orr, 2003), little has been studied about the detailed cognitive manifestations mediated by

environmental factors, especially in the psycholinguistic domain. The similarities between low SES poor readers and dyslexics suggested in the current study could be an example to provide scientific based knowledge for developing educational and interventional strategies for the low SES readers based on the dyslexia literature. Nevertheless, more awareness about environmental enrichment needs to be raised in research, clinical, educational, and legislative endeavors, in order to develop and apply the optimal educational and interventional strategies to the low SES poor readers.

To conclude, our measures of the fine-grained time course of lexical activation during spoken word recognition in response to misleading coarticulation suggest that poorer reading abilities are associated with higher sensitivity to subcategorical phonetic detail, consistent with the over-specification hypothesis (Serniclaes, 2006). Our data collected from low SES young adults suggested that poor environment could have influential and detrimental effects on individuals' phonological representations and reading abilities. Therefore, our findings merit attention from educational and interventional perspectives to address this long-neglected public health issue. Nevertheless, more computational modeling work and empirical evidence are required to further verify our arguments.

Appendix 1

Target (W1)	Word Competitor (W2)	Non-word Competitor (N3)
SIMILAR		
bat	bag	bab
bud	bug	bub
butt	buck	bup
fort	fork	forp
hood	hook	hoop
net	neck	nep
pit	pig	pib
rod	rock	rop
tap	tack	tat
DISSIMILAR		
beak	bead	beab
carp	cart	cark
cat	cab	cag
harp	heart	hark
knot	knob	knog
road	rope	roke

Adapted from Appendix A of Dahan et al. (2001).

Note. Stimulus triplets were categorized based on the similarity of final consonants' place of articulation between W1 and N3. Similar: the final consonants of W1 and N3 were either labial or alveolar; dissimilar: one of the final consonants of W1 and N3 was velar, and the other was either labial or alveolar.

Appendix 2

Target (W1)	Competitor (W2)	Distractor 1	Distractor 2
bat	bag	pen	stool
beak	bead	saw	thumb
bud	bug	fox	eye
butt	buck	clams	ghost
carp	cart	swing	moon
cat	cab	vase	tree
fort	fork	light	hat
harp	heart	desk	claw
hood	hook	eggs	brush
knot	knob	mouse	beer
net	neck	bass	deer
pit	pig	ark	flute
road	rope	knee	glass
rod	rock	bear	fries
tap	tack	skunk	peas

Adapted from Appendix B of Dahan et al. (2001).

References

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*, *100*(2), 235–251. http://doi.org/10.1037/0022-0663.100.2.235

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, *38*(4), 419–439. http://doi.org/10.1006/jmla.1997.2558

Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition*, *111*(1), 55–71. http://doi.org/10.1016/j.cognition.2008.12.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., & Sprenger-Charolles, L. (2008). Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *Journal of Experimental Child Psychology*, *101*(2), 137–155. http://doi.org/10.1016/j.jecp.2008.03.006

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252.

Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., … Shankweiler, D. P. (2016). Vocabulary does not complicate the simple view of reading. *Reading and Writing*, *29*(3), 435–451. http://doi.org/10.1007/s11145-015-9608-6

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226–243. http://doi.org/10.1177/00222194070400030401

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology*, *28*(5), 874–886. http://doi.org/10.1037/0012-1649.28.5.874

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(3), 687–696. http://doi.org/10.1037/0278-7393.30.3.687

Coley, R. J. (2002). An uneven start: Indicators of inequality in school readiness. Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/research/pic

Corkin, S. (1974). Serial-ordering deficits in inferior readers. *Neuropsychologia*, *12*(3), 347–354. http://doi.org/10.1016/0028-3932(74)90050-5

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, *16*(5–6), 507–534. http://doi.org/10.1080/01690960143000074

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test (3rd ed.). Circle Pines, MN: American Guidance Service, Inc.

Elbro, C. (1998). When reading is "readn" or somthn. Distinctness of phonological representations of lexical items in normal and disabled readers. *Scandinavian Journal of Psychology*, *39*(3), 149–153. http://doi.org/10.1111/1467-9450.393070

Elbro, C., Borstrøm, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: the importance of distinctiveness of phonological represetations of lexical items. *Reading Research Quarterly*, *33*(1), 36–60. http://doi.org/10.1598/RRQ.33.1.3

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language*, *54*(4), 554–573. http://doi.org/10.1016/j.jml.2005.12.009

Fischer, B. (1992). Saccadic Reaction Time: Implications for Reading, Dyslexia, and Visual Cognition. In *Eye movements and visual cognition* (pp. 31–45). Springer New York. http://doi.org/10.1007/978-1-4612-2852-3_3

Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression* (Second). Thousand Oaks CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*, *106*(3), 491–528. http://doi.org/10.1037/0033-295X.106.3.491

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662–720. http://doi.org/10.1037/0033-295x.111.3.662

Hollingshead, A. B. (1975). Four factor index of social status. New Haven, CT: Yale University, Department of Sociology.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), B23-32. http://doi.org/10.1016/j.cognition.2004.10.003

Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2009). Structural and semantic

constraints on the resolution of pronouns and reflexives. *Cognition*, *112*(1), 55–80. http://doi.org/10.1016/j.cognition.2009.03.010

Karlson, B., & Gardner, E. (1995). Stanford Diagnostic Reading Test (4th ed.). San Antonio, TX: Psychological Corporation.

Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., … Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica*, *171*, 72–84. http://doi.org/10.1016/j.actpsy.2016.09.009

Lagae, L. (2008). Learning Disabilities: Definitions, Epidemiology, Diagnosis, and Intervention Strategies. *Pediatric Clinics of North America*, *55*(6), 1259–1268. http://doi.org/10.1016/j.pcl.2008.08.001

Liberman, I. Y. (1973). Segmentation of the spoken word and reading acquisition. *Bulletin of the Orton Society*, *23*(1), 64–77. http://doi.org/10.1007/BF02653842

Luce, P. A. (1986). *Neighborhoods of Words in the Mental Lexicon. Research on Speech Perception. Technical Report No. 6. Doctoral Dissertations*. Indiana University.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). Gates-MacGinitie Reading Tests Level AR: Directions for Administration. Itasca, IL: Riverside Publishing.

Magnuson, J. S., Kukona, A., Braze, D., Johns, C. L., Van Dyke, J. A., Tabor, W., … Shankweiler, D. P. (2011). Phonological instability in young adult poor readers: Time course measures and computational modeling. In P. McCardle, B. Miller, J. R. Lee, & O. Tzeng (Eds.), *Dyslexia Across Languages: Orthography and the Brain-Gene-Behavior Link* (pp. 184–201). Baltimore, Maryland: Paul H. Brookes.

Magnuson, J. S., & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the

perceptual accommodation of talker variability. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(2), 391–409. http://doi.org/10.1037/0096-1523.33.2.391

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: studies with artificial lexicons. *Journal of Experimental Psychology. General*, *132*(2), 202–227. http://doi.org/10.1037/0096-3445.132.2.202

Markwardt, F. C. (1989). Peabody Individual Achievement Test - Revised. Circle Pines, MN: American Guidance Service, Inc.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*(4), 653–675. http://doi.org/10.1037/0033-295X.101.4.653

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. http://doi.org/10.1016/0010-0285(86)90015-0

McDougall, S., Hulme, C., Ellis, a, & Monk, a. (1994). Learning to read: the role of short-term memory and phonological skills. *Journal of Experimental Child Psychology*. http://doi.org/10.1006/jecp.1994.1028

McMurray, B., Aslin, R. N., Tanenhaus, M. K., Spivey, M. J., & Subik, D. (2008). Gradient sensitivity to within-category variation in words and syllables. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(6), 1609–1631. http://doi.org/10.1037/a0011747

McMurray, B., Munson, C., & Tomblin, J. B. (2014). Individual Differences in Language Ability Are Related to Variation in Word Recognition, Not Speech Perception: Evidence From Eye Movements. *Journal of Speech Language and Hearing Research*, *57*(4), 1344.

http://doi.org/10.1044/2014_JSLHR-L-13-0196

McQueen, J. M., & Cutler, A. (2001). Spoken word access processes: An introduction. *Language and Cognitive Processes*, *16*(August 2011), 469–490. http://doi.org/10.1080/01690960143000209

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. Chapman & Hall/CRC.

Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving Schools in Socio-economically Disadvantaged Areas: a Review of Research Evidence. *School Effectiveness and School Improvement*, *15*(April), 37–41. http://doi.org/10.1076/sesi.15.2.149.30433

Muñoz, D. P., Everling, S., Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*(3), 218–228. http://doi.org/10.1038/nrn1345

Noordenbos, M. W. (2013). *Phonological representations in dyslexia: Underspecified or overspecified? (Doctoral dissertation)*. Radboud University Nijmegen, The Netherlands. Retrieved from http://repository.ubn.ru.nl/bitstream/handle/2066/115718/115718.pdf?sequence=1

Orr, A. J. (2003). Black-White Differences in Achievement: The Importance of Wealth. *Sociology of Education*, *76*(4), 281–304. http://doi.org/10.2307/1519867

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Manual, Vienna, Austria. Retrieved from https://www.r-project.org/

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, *13*(2), 212–218. http://doi.org/10.1016/S0959-4388(03)00035-7

Ramus, F., Marshall, C. R., Rosen, S., & Van Der Lely, H. K. J. (2013). Phonological deficits in

specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, *136*(2), 630–645. http://doi.org/10.1093/brain/aws356

Serniclaes, W. (2006). Allophonic perception in developmental dyslexia: Origin, reliability and implications of the categorical perception deficit. *Written Language & Literacy*, *9*(1), 135–152. http://doi.org/10.1075/wll.9.1.09ser

Serniclaes, W., Sprenger-Charolles, L., Carré, R., & Démonet, J.-F. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language, and Hearing Research : JSLHR*, *44*(2), 384–399. http://doi.org/10.1044/1092-4388(2001/032)

Serniclaes, W., Van Heghe, S., Mousty, P., Carré, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology*, *87*(4), 336–361. http://doi.org/10.1016/j.jecp.2004.02.001

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 531–545. http://doi.org/10.1037/0278-7393.5.6.531

Shankweiler, D., Mencl, W. E., Braze, D., Tabor, W., Pugh, K. R., & Fulbright, R. K. (2008). Reading Differences and Brain: Cortical Integration of Speech and Print in Sentence Processing Varies With Reader Skill. *Developmental Neuropsychology*, *33*(6), 745–775. http://doi.org/10.1080/87565640802418688

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family Risk of Dyslexia Is Continuous: Individual Differences in the Precursors of Reading Skill. *Child Development*, *74*(2), 358–373. Retrieved from http://www.jstor.org/stable/3696318

Snowling, M. J., & Hayiou-Thomas, M. E. (2006). The Dyslexia Spectrum. *Topics in Language*

Disorders, *26*(2), 110–126. http://doi.org/10.1097/00011363-200604000-00004

Sprenger-Charolles, L., Colé, P., Lacert, P., & Serniclaes, W. (2000). On subtypes of developmental dyslexia: evidence from processing time and accuracy scores. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, *54*(2), 87–104. http://doi.org/10.1037/h0087332

Sprenger-Charolles, L., Siegel, L. S., Béchennec, D., & Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, *84*(3), 194–217. http://doi.org/10.1016/S0022-0965(03)00024-9

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: the cognitive correlates of print exposure. *Memory and Cognition*, *20*(1), 51–68. http://doi.org/10.3758/BF03208254

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

Tallal, P., Merzenich, M. M., Miller, S., & Jenkins, W. (1998). Language learning impairments: Integrating basic science, technology, and remediation. *Experimental Brain Research*, *123*(1–2), 210–219. http://doi.org/10.1007/s002210050563

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. http://doi.org/10.1126/science.7777863

Torgeson, J. K., Wagner, R. K., & Rashotte, C. A. (1999). Tests of Word Reading Efficiency (TOWRE). Austin, TX: Pro-Ed.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). {mice}: Multivariate Imputation by Chained

Equations in R. *Journal of Statistical Software*, *45*(3), 1–67. article. Retrieved from http://www.jstatsoft.org/v45/i03/

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth). New York: Springer.

Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual, and motor control aspects. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 353–393). Amsterdam: Elsevier.

Wagner, R. K., Torgesen, J., & Rashotte, C. (1999). The comprehensive test of phonological processing (CTOPP). Austin, TX: Pro-Ed.

Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. New York, NY: The Psychological Corporation: Harcourt Brace & Company.

Wiederholt, J. L., & Bryant, B. R. (2001). Gray Oral Reading Test (GORT) (4th ed.). Austin, TX: Pro-Ed.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III Tests of Achievement. Itasca, IL: Riverside Publishing.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. http://doi.org/10.1037/0033-2909.131.1.3