SCIENTIFIC REP{O}RTS

OPEN

Reading ability and phoneme categorization

Gabrielle E. O'Brien, Daniel R. McCloy ⓘ, Emily C. Kubota & Jason D. Yeatman

Dyslexia is associated with abnormal performance on many auditory psychophysics tasks, particularly those involving the categorization of speech sounds. However, it is debated whether those apparent auditory deficits arise from (a) reduced sensitivity to particular acoustic cues, (b) the difficulty of experimental tasks, or (c) unmodeled lapses of attention. Here we investigate the relationship between phoneme categorization and reading ability, with special attention to the nature of the cue encoding the phoneme contrast (static versus dynamic), differences in task paradigm difficulty, and methodological details of psychometric model fitting. We find a robust relationship between reading ability and categorization performance, show that task difficulty cannot fully explain that relationship, and provide evidence that the deficit is not restricted to dynamic cue contrasts, contrary to prior reports. Finally, we demonstrate that improved modeling of behavioral responses suggests that performance does differ between children with dyslexia and typical readers, but that the difference may be smaller than previously reported.

Dyslexia is a learning disability that affects between 5% and 17% of the population and poses a substantial economic and psychological burden for those affected[1–3]. Despite decades of research, it remains unclear why so many children without obvious intellectual, sensory, or circumstantial challenges find written word recognition so difficult.

One popular and persistent theory is that dyslexia arises as a result of an underlying auditory processing deficit[4–9]. According to this theory, a low-level auditory processing deficit disrupts the formation of a child's internal model of speech sounds (phonemes) during early language learning; later, when young learners attempt to associate written letters (graphemes) with phonemes, they struggle because their internal representation of phonemes is compromised[10].

In line with this hypothesis, many studies report group differences between dyslexics and typical reading control participants in auditory psychophysical tasks including amplitude modulation detection[11–15], frequency modulation detection[16–20], rise time discrimination, and duration discrimination[21–25]. Moreover, attributing dyslexia to an auditory deficit is appealing because one of the most effective predictors of reading difficulty is poor phonemic awareness—the ability to identify, segment and manipulate the phonemes within a spoken word[26,27]. If a child's phoneme representation were abnormal due to auditory processing deficits, it could be a common factor underlying both poor performance in auditory phoneme awareness tests and difficulties learning to decode written words.

In one variant of this auditory hypothesis, dyslexia is thought to involve a deficit specifically in the processing of rapid modulations in sound, usually referred to as a "rapid temporal processing" deficit[5,28,29]. This idea has been controversial (see Rosen[30,31] for review), but also remains as one of the widely cited accounts of auditory deficits in dyslexia[7]. One source of controversy is that the rapid temporal processing deficit is far from universal; in Tallal's[5] study that first proposed a causal relationship between rapid temporal processing and reading ability, only 8 out of 20 dyslexic children showed a deficit on a temporal order judgment task. Similarly, in a more comprehensive study of 17 dyslexic adults, only 11 had impaired performance on an extensive battery of auditory psychophysical tasks[32]. Moreover, in that study, tasks requiring temporal cue sensitivity were not systematically more effective than non-temporal tasks at separating dyslexics from controls. Several studies have also shown that dyslexic participants have heightened sensitivity to allophonic speech contrasts (differences in pronunciation that do not change the identity of the speech sound)[33–35], including contrasts primarily marked by temporal differences in the stimuli, which contradicts the hypothesis that the dyslexic brain lacks access to temporal information. Furthermore, other studies have shown that dyslexics have normal abilities to resolve spectrotemporal

Institute for Learning and Brain Sciences and Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA. Correspondence and requests for materials should be addressed to G.E.O. (email: eobrien3@uw.edu)

1

modulations in noise, suggesting that apparent speech perception difficulties result from higher-level aspects of processing beyond encoding sounds[36–38]. More recently, the rapid temporal processing hypothesis has been reframed as a deficit in processing dynamic, but not necessarily rapid, aspects of speech[10,39,40].

On the whole, there is ample evidence that abnormal performance on auditory psychophysical tasks is frequently associated with dyslexia[41], but much less evidence that a specific deficit in temporal processing is causally related to dyslexia. But regardless of whether the alleged deficit involves all modulations of the speech signal or only rapid ones, there are at least two major reasons why the theory that auditory temporal processing deficits cause the reading difficulties seen in dyslexia has been called into question. First, although correlations have been found between certain psychoacoustic measures and reading skills[14,42,43], there have been no direct findings relating performance on psychoacoustic tasks to phonological representations[31]. On the contrary, there is growing evidence that poor performance on psychophysical tasks may be partially or totally explained by the demands of the test conditions and not the stimuli themselves[25]. In this view, it is not fuzzy phonological representations, but rather poor working memory—which often co-occurs with dyslexia[44–47]—that drives deficits in both psychoacoustic task performance and in reading ability[24,48]. Proponents of this hypothesis argue that the phonological deficit associated with dyslexia is only observed when considerable memory and time constraints are imposed by the task[49]; evidence for this view comes from studies in which listeners perform different kinds of discrimination paradigms on the same auditory stimuli[25,50–52]. A neuroimaging study indicating that dyslexics have similar neural representations of speech sounds as controls, but differences in connectivity between speech processing centers, also lends support to the idea of "intact but less accessible" representations of phonetic features[53]. Additionally, many dyslexics perform normally on speech perception tasks when provided ideal listening conditions but exhibit a deficit when speech is presented in noise. This finding has been interpreted as indicating a deficit in higher-level auditory processing such as the ability to attend to perceptually relevant cues and ignore irrelevant aspects of the signal[36,37,52]. Dyslexia is also known to be comorbid with ADHD, and numerous reviews have suggested that, in addition to working memory, differences in attention may be another key driver of the observed group differences[32,54]. However, it is not clear how this hypothesis accounts for several studies where dyslexics actually perform *better* than controls discriminating fine-grained acoustic cues[22,34,36].

A second reason to question the link between auditory processing deficits and dyslexia is that the standard technique for data analysis in psychophysical experiments is prone to severe bias. Specifically, in many phoneme categorization studies, participants identify auditory stimuli that vary along some continuum as belonging to one of two groups, a psychometric function is fit to the resulting data, and the slopes of these psychometric functions are compared between groups of dyslexic and control readers (for reviews, see Vandermosten et al.[55]; Noordenbos and Serniclaes[35]). In this approach, a steep slope indicates a clear boundary between phoneme categories, whereas a shallow slope suggests less defined categories (i.e., fuzzy phonological representations), possibly due to poor sensitivity to the auditory cue(s) that mark the phonemic contrast. Unfortunately, most studies in the dyslexia literature fit the psychometric function using algorithms that fix its asymptotes at zero and one, which is equivalent to assuming that categorization performance is perfect at the extremes of the stimulus continuum (i.e., assuming a "lapse rate" of zero). This assumption is questionable in light of the evidence that dyslexics may be less consistent categorizing stimuli across the full continuum[35], as well as evidence that attention, working memory, or task difficulty, rather than stimulus properties, may underlie group differences between readers with dyslexia and control subjects.

The zero lapse rate assumption is particularly problematic given that fixed asymptotes at zero and one leads to strongly downward-biased slope estimates even when true lapse rates are fairly small[56,57]. In other words, if a participant makes chance errors on categorization, and these errors happen on trials near the ends of the continuum, the canonical psychometric fitting routine used in most previous studies will underestimate the steepness of the category boundary. Thus, a tendency to make a larger number of random errors (due to inattention, memory, or task-related factors) will be wrongly attributed to an indistinct category boundary on the contrast under study. Since it is precisely the subjects with dyslexia that more frequently show attention or working memory deficits, research purporting to show less distinct phoneme boundaries in readers with dyslexia may in fact reflect non-auditory, non-linguistic differences between study populations.

Unfortunately, most studies in the dyslexia literature that use psychometric functions to model categorization performance appear to suffer from this bias. Although most do not report their analysis methods in sufficient detail to be certain, we infer (based on published plots, and the lack of any mention of asymptote estimation) that the bias is widespread (e.g., Reed[58]; Manis et al.[59]; Breier et al.[60]; Chiappe, Chiappe and Siegel[61]; Maassen et al.[62]; Bogliotti et al.[34]; Zhang et al.[63]). A few studies report using software that in principle supports asymptote estimation during psychometric fitting, but do not report the parameters used in their analysis (e.g., Vandermosten et al.[43,55]). In one case, researchers who re-fit their psychometric curves without data from the continuum endpoints found a reduced effect size, prompting the authors to wonder whether any effect would remain if an unbiased estimator of slope were used[64]. Only a few studies have fit psychometric functions with free asymptotic parameters or investigated differences in lapse rates[35,65,66]; however, it is worth noting that their methods did not constrain the lapse rate, which can lead to upward-biased slope estimates[67]. This is a consequence of the fact that slope and lapse parameters "trade off" in the optimization space of sigmoidal function fits[68].

In light of all these problems—the inconsistency of findings, confounding influences of experimental task design, and bias introduced by analysis choices—it is reasonable to wonder whether children with dyslexia really have any abnormality in phoneme categorization, once those factors are all controlled for. The present study addresses the relationship between reading ability and phoneme categorization ability, and in particular, whether children with dyslexia show a greater deficit on phoneme contrasts that rely on temporally varying (dynamic) cues, as opposed to static cues. This study avoids the aforementioned methodological problems by using multiple paradigms with different attentional and memory demands, and by analyzing categorization performance using Bayesian estimation of the psychometric function[67]. In this approach, the four parameters of a psychometric
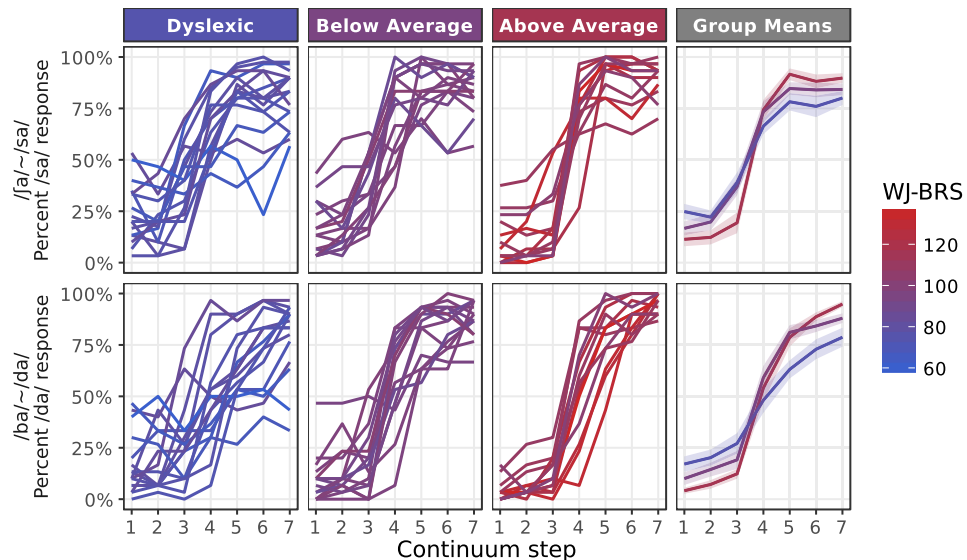
Figure 1. Response functions for each subject and each speech continuum. Results are averaged over responses from both the ABX and single-interval paradigms. The color of each line maps to the subject's reading ability based on the Woodcock Johnson Basic Reading Skills Composite score (WJ-BRS).

function—the threshold, slope, and two asymptote parameters—are assigned prior probability distributions, which formalize the experimenter's assumptions about their likely values. By allowing the asymptote parameters to vary, the slope is estimated in a less biased way than in traditional modeling approaches with fixed or unconstrained asymptotes. However, because fitting routines trade off between asymptote parameters and the slope parameter of a logistic model in optimization space[68], it can be difficult to estimate both accurately at the same time. To address this difficulty, we first performed cross-validation on the prior distribution of the asymptote parameters to determine the optimal model to fit the data.

This paper presents data from 44 children, aged 8–12 years, and a wide range of reading abilities. Our experimental task is based on the design of Vandermosten et al.[43,55], which assessed categorization performance for two kinds of stimulus continua: those that differed based on a spectrotemporal cue (dynamic), and those that differed based on a purely spectral cue (static). In the original study, the authors concluded that children with dyslexia are specifically impaired at categorizing sounds (both speech and non-speech) that differ on the basis of dynamic cues. However, although the dynamic and static stimuli in their study were equated for overall length, the duration of the cues relevant for categorization were not equal: in the dynamic stimuli, the cue (a vowel formant transition) was available for 100 ms, but in the static stimuli, the cue (the formant frequency of a steady state vowel) was available for 350 ms. This raises the question of whether cue duration, rather than the dynamic nature of the cue, was the source of apparent impairment in categorization among participants with dyslexia. The present study avoids this confound by changing the "static cue" stimuli from steady-state vowels to fricative consonants (a /ʃa/~/sa/continuum), so that the relevant cue duration is 100 ms in both the static (/ʃa/~/sa/) and dynamic (/ba/~/da/) stimulus continua. Additionally, Vandermosten and colleagues used a test paradigm in which listeners heard three sounds and were asked to decide if the third sound was more like the first or second (an ABX design). Here we included both an ABX task and a single-stimulus categorization task, to see whether the memory and attention demands of the ABX paradigm may have played a role in previous findings. Thus, by (a) assessing categorical perception of speech continua with static and dynamic cues, (b) varying the cognitive demands of the psychophysical paradigm, and (c) empirically determining the optimal parameterization of the psychometric function with cross-validation, we aim to clarify the role of auditory processing deficits in dyslexia.

Results

Response functions for each subject on the two stimulus continua (aggregated across task paradigms) are shown in Fig. 1. Subjects were divided into three groups (Dyslexic, Below Average, and Above Average readers) to be consistent with previous studies that have reported group comparisons, and reading score (as indexed by the Woodcock Johnson Basic Reading Skills standard score (WJ-BRS)) was also treated as a continuous variable in line with the perspective that dyslexia represent the lower end of a continuous distribution[69].

For Above Average readers, response functions show the typical sigmoid shape, with a steep slope marking the boundary between categories and consistent labeling within categories; for readers in the Below Average and Dyslexic groups, response functions are much more heterogeneous, with some subjects exhibiting inconsistent performance even at continuum endpoints. The group means are suggestive of differences in both psychometric slope and lapse rate on both the static (/ʃa/~/sa/) and dynamic (/ba/~/da/) cue continua.

Figure 2 shows the psychometric functions fit to each subject's response data on the two continua (aggregated across paradigms). The fitted curves reflect the same pattern seen in the raw data: subjects with high reading scores tend to show consistent categorization at stimulus endpoints and steep category transitions near the center
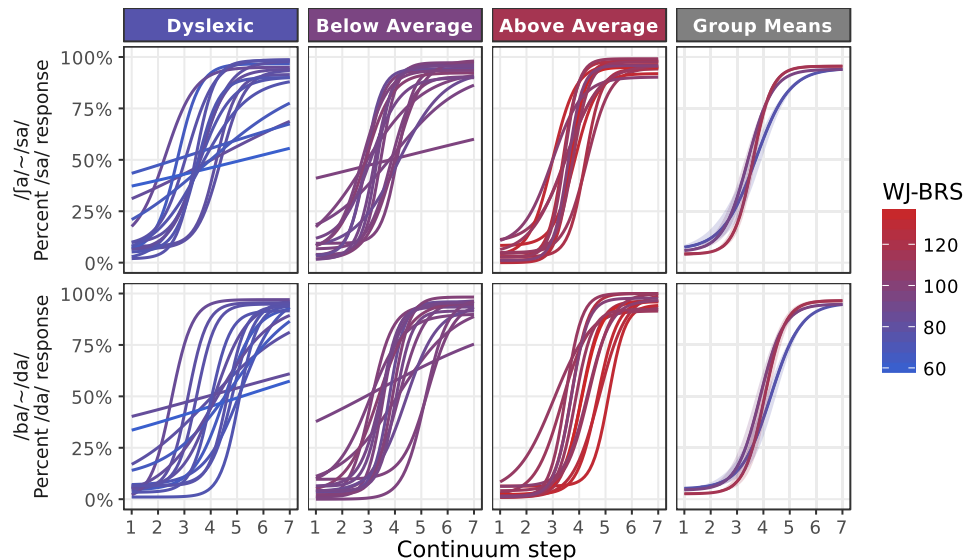
Figure 2. Fitted response functions for each subject and each continuum. Parameters (slope, threshold, and the two asymptotic parameters) were averaged over estimates from both the ABX and single-interval paradigms. The color of each line maps to the subject's reading ability (WJ-BRS score).

of the stimulus continuum, whereas subjects with poor reading scores show a diversity of curve shapes. For some subjects with poor reading scores the psychometric curves appear quite similar to those of the skilled readers, while others exhibit shallow slopes at the category transition and/or high lapse rates at the continuum endpoints. Estimates of the category boundary (the "threshold" parameter of the psychometric) also appear to span a wider range of continuum values among poor readers compared with readers in the above average group.

High reading score predicts consistent phoneme identification. The group-level patterns of slope and lapse rate estimates are summarized in Fig. 3. There is a clear trend of increasingly steeper slopes for children with higher reading scores, across both stimulus continua and both experimental paradigms. The pattern of lapse rate estimates is less consistent, though the lapse rate estimates tended to be higher for poor readers. Slope and lapse rate estimates were each modeled using mixed effects regression, as implemented in the lme4 library for R[70]. Fixed-effect predictors with deviation coding were used for both the continuum (static /ʃa/~/sa/versus dynamic/ ba/~/da/) and paradigm (ABX versus single-stimulus) variables, and reading ability (WJ-BRS score) was included as a continuous fixed-effect predictor. A random intercept by participant was also included. Fixed-effect predictors were also included to control for the effects of ADHD diagnosis and nonverbal IQ (age-normalized WASI Matrix Reasoning score). Because two blocks of the ABX test were administered per subject, slope or lapse rate parameter estimates were averaged over the two blocks for each individual before regression modeling.

The fully-specified model of psychometric slope estimates indicated that only one predictor—WJ-BRS reading score—had a significant effect on slope estimates; there was no reliable difference in psychometric slopes between the static and dynamic cue continua, nor was there a reliable difference between ABX blocks and single-interval blocks. Stepwise nested model comparisons were then used to eliminate irrelevant model parameters and yield the most parsimonious model (see supplement for details). The most parsimonious model of psychometric slope contained only a continuous predictor of reading ability (WJ-BRS) and a random intercept for each participant (see Table 1). The effect size of reading ability on psychometric slope was relatively modest; a 1-unit improvement in WJ-BRS score was associated with an increase of 0.023 in the slope of the psychometric function (slope is measured in units of probability (of selecting /da/ or /sa/) per step along the 7-step continua). This indicates that poor readers tend to have shallower psychometric slopes (p = 0.0005). For comparison, increasing age by 1 year was associated with an increase of 0.256 in the slope of the psychometric function. In other words, it takes an 11-point increase in reading score to approximate the change in slope associated with a 1-year difference in age.

To model the relationship between lapse rate and reading score, we first combined the two asymptote estimates from the psychometric fits, by averaging the asymptote deviation from zero or one (i.e., a lower asymptote of 0.1 and an upper asymptote of 0.9 both correspond to a lapse rate of 0.1; the upper and lower lapse rates were averaged for each psychometric fit). The same initial model specification and simplification procedure used with the slope model was also used for the lapse rate model. The final model of lapse rate contained a continuous predictor for reading ability, a categorical predictor for continuum, and a random effect for participant (see Table 1); no reliable effect of task paradigm was detected. Higher reading ability was associated with smaller lapse rates (about 0.04% smaller per 1-unit increase of WJ-BRS score), and the static /ʃa/~/sa/ continuum was associated with larger lapse rates (about 1% higher than the lapse rate estimates for the /ba/~/da/ continuum). Thus, compared to strong readers, poor readers show shallower psychometric slopes and are also more likely to make errors even near the continuum endpoints. There was a general tendency for all participants to make more errors near continuum endpoints with the static-cue stimuli than with the dynamic-cue stimuli.
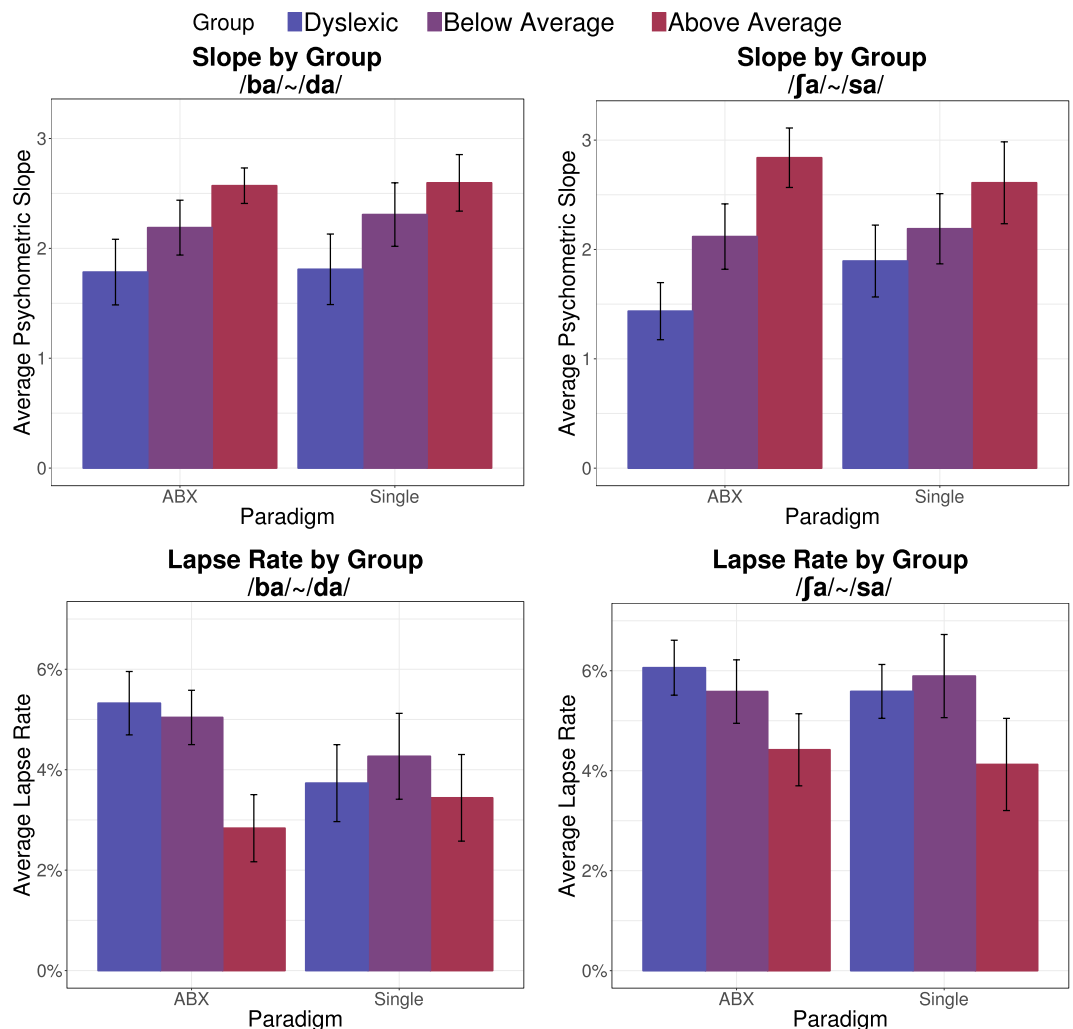
Figure 3. Average psychometric slope within groups for the /ba/~/da/ and /ʃa/~/sa/ continua are shown in panels A and B, separated by paradigm. Average lapse rate parameters for both continua are shown in Panels C and D. Error bars represent one standard error of the mean.

Model	Parameter	β	SE	p
4-parameter slope	WJ-BRS	0.023	0.006	0.0005
4-parameter lapse	WJ-BRS	−0.0004	0.0001	0.002
	Continuum (static)	0.010	0.003	0.007
2-parameter slope	WJ-BRS	0.017	0.004	0.0004
	Continuum (static)	−0.167	0.078	0.039
1st principal component	WJ-BRS	0.021	0.006	0.0007
	WJ-BRS:Paradigm	0.012	0.006	0.048

Table 1. Summary of significant model parameters for the four linear mixed effects models.

We next analyzed these relationships treating reading level as a categorical rather than a continuous variable. Using a mixed model analysis with group (Dyslexic, Below Average, and Above Average) as a categorical predictor with Above Average as the reference group, and subject as a random factor, and following the same model selection procedure as before, we confirmed the relationship between group and psychometric slope ($F_{(2,40.72)} = 6.614$, $p = 0.003$). The selected model did not include effects of continuum or paradigm. To assess the separability of the groups, we calculated Cohen's d for a group comparison where each individual's slope estimate is averaged across the six test blocks (combining paradigms and continua). We found a large effect size for the comparison between the Dyslexic and Above Average groups ($d = 1.48$) and a medium effect size for the comparison between Dyslexic and Below Average groups ($d = 0.66$). Indeed a linear discriminant analysis (LDA) classification model was able to accurately classify subjects as Dyslexic versus Above Average readers 72.4% of the

time (prediction accuracy assed using leave-one-out cross validation[71]). The same model achieved 51.5% accuracy when classifying Dyslexics versus Below Average readers.

An analysis of lapse rates across reading groups found a significant main effect of group ($F(2,41.09) = 5.298$, $p = 0.009$). Belonging to the Below Average group was associated with an increased lapse rate by 1.48% ($p = 0.0140$) and belonging to the Dyslexic group was associated with an increased lapse rate by 1.76% ($p = 0.004$) compared to the Above Average group. Using a quadratic discriminant analysis (QDA) with both slope and lapse rate did not improve cross-validated prediction accuracy for classifying Dyslexics versus Above Average readers (72.4%) but did improve the ability to classify Dyslexics versus Below Average readers (64.5%).

Traditional model fitting misrepresents effect of cue continuum.

For comparison with the methods commonly used in prior studies, we also fit a mixed effects model using slope estimates from a traditional 2-parameter psychometric function fit (with asymptotes fixed at 0 and 1, see Methods), and performed the same model simplification procedure. In this case, the most parsimonious model indicated a significant effect of stimulus continuum on psychometric slope, in addition to the effect of reading ability seen in the model of 4-parameter fits (see Table 1). The magnitude of the estimated effect of continuum on psychometric slope is quite large (difference in slope of 0.167 between dynamic and static cue continua, comparable to a difference of about 10 points on the WJ-BRS), with lower slopes associated with the static-cue (/ʃa/~/sa/) continuum. Comparing this to the models of 4-parameter fits, we see reasonably similar estimates of the effect of reading ability on psychometric slope, but the effect of stimulus continuum is allocated differently depending on the underlying assumptions of the psychometric fitting routine.

Principal component analysis reveals an effect of task paradigm.

Although slope and lapse rate are often thought of as representing separate cognitive processes (category discrimination and attention, respectively), in reality it is difficult to disentangle what is being modeled by each parameter: the two are not truly independent in the psychometric fitting procedure. Moreover, in speech continua, slope and lapse rate may be influenced by the same underlying processes: the endpoints are determined by natural speech tokens, which may not be reliably categorized 100% of the time by all subjects, due to the complex relationships among speech cues, talker/listener dialect, context, etc. Abnormal categorical perception that manifests as shallower psychometric slope may also affect the way stimuli at the endpoints of the continuum are labelled[33]. For these reasons, we also performed a principal components analysis that combined slope and asymptote estimates into a single index of psychometric curve shape. The first principal component was a linear combination of the slope and both asymptotes (slope: 0.678; lower asymptote: -0.268; upper asymptote: -0.685) that explains about 41% of variance in the data. Using an identical modeling procedure to the analysis of slope and lapse rate, we iteratively eliminated predictors from a fully specified mixed model of the first principal component. The most parsimonious model showed a main effect of reading ability (WJ-BRS score) and a significant interaction between reading ability and experimental paradigm, but no effect of stimulus continuum (see Table 1). The main effect indicated that poor readers tend toward a combination of higher lapse rates and shallower slopes, while strong readers tend to have lower lapse rates and steeper slopes. The once in the ABX paradigm than in the single-interval paradigm. Thus, in line with previous hypotheses[25,49], this model suggests that the specific paradigm used in a psychophysical experiment can amplify deficits in people with dyslexia. However, struggling readers still perform poorly on phoneme categorization tasks irrespective of the paradigm used in the experiment.

Poor readers struggle even with unambiguous stimuli at continuum endpoints.

The lapse rate results discussed above suggest that children with dyslexia may have considerable difficulty even in categorizing unambiguous stimuli at the continuum endpoints. We performed post-hoc analyses of accuracy and reaction time to further investigate this possibility. In particular, we were interested in whether the high lapse rate estimates resulted from poor readers' performance degrading over the course of an experimental block, indicating an inability to stay focused and pay attention for the duration of the experiment. We were also interested in whether the pattern of reaction times across different parts of the stimulus continuum varied systematically with reading ability.

Accuracy at continuum endpoints. Because our young subjects ranged considerably in attention span and motivation, we analyzed the stationarity of response accuracy at continuum endpoints, to ascertain whether task performance declined from start to finish. Following the design of Messaoud-Galusi *et al.*[64], for each test block (6 total per subject) we computed the number of correct assignments of the stimulus endpoints. There were twenty endpoint stimulus presentations per block.

We then modeled correct endpoint stimulus categorization as a function of trial number (1 through 70), reading score, stimulus continuum, and test paradigm (ABX or Single-stimulus), plus covariates for age and ADHD diagnosis and a random effect of subject. WJ-BRS, age, and trial number were centered for clear interpretation of interaction terms. Following the same model fitting and simplification procedures described above, we found the most parsimonious model included only main effects of WJ-BRS ($\beta = 0.004$, SE $= 0.001$, $p = 0.001$) and trial number ($\beta = -0.0008$, SE $= 0.0003$, $p = 0.022$). The main effect of WJ-BRS indicated that poor readers were more likely than typical readers to err when categorizing unambiguous stimuli, echoing the earlier analysis of lapse rate. The main effect of trial number indicates that all readers tended to decline in accuracy as the test block progressed. The lack of a significant interaction between trial number and WJ-BRS indicates that on average, poor readers were no more likely than other subjects to decline in performance due to fatigue or inattention as the test block wore on. Thus, inability to stay focused over the course of a psychophysics experiment does not explain phoneme categorization deficits in our sample of poor readers.

Reaction time analysis. With stimulus continua that span a categorical boundary, it is expected that reaction times will be longer for stimuli near the category boundary[72,73], reflecting the perceptual uncertainty and resulting decision-making difficulty associated with ambiguous stimuli. Although our participants were not under time constraints in this study, they were encouraged to respond as soon as they could after stimulus presentation, and reaction times were recorded. To see whether the children we tested showed the expected pattern (faster reaction times on unambiguous stimuli at or near continuum endpoints, and slower reaction times on ambiguous stimuli near continuum centers), we modeled reaction times across continuum steps with a second-order polynomial basis (and its interaction with reading ability).

Kernel density estimates of the reaction time distributions for the three groups are shown in Fig. 4A, individual and group-level polynomial fits are shown in Fig. 4B,C, and the quadratic coefficients of the polynomial fits are shown in Fig. 4D. The figure separates data from the two stimulus continua for easier comparison with Figs 1 and 2, but since we had no specific hypothesis about reaction time differences between continua, we did not include that comparison in the post-hoc analysis.

Model results showed an overall relationship between reading ability and reaction time ($\beta = -0.0034$, $SE = 0.0009$, $p = 1.69e\text{-}4$) indicating an average 2.5 ms speed-up in reaction time per unit change in WJ-BRS score, although there was substantial overlap in reaction time distributions across groups (cf. Fig. 4A). The quadratic polynomial coefficient was also significant ($\beta = -2.58$, $SE = 0.447$, $p = 7.83e\text{-}9$) showing a general tendency for faster responses at continuum endpoints. Finally, there was a significant interaction between the quadratic polynomial coefficient and WJ-BRS score ($\beta = -0.082$, $SE = 0.024$, $p = 5.29e\text{-}4$), indicating a more downward-curving polynomial for subjects with higher reading scores (cf. Fig. 4B–D). In other words, good reading ability was associated with the expected pattern of reaction times along the stimulus continuum, whereas poor readers were less likely to show an expected pattern; many individual fits showed nearly flat fitted curves or even an inverse pattern to that of the above-average readers. This analysis indicates that children with dyslexia may experience synthesized speech continua in a qualitatively different way from typical readers. Specifically, children with dyslexia are more likely to respond endpoint stimuli and ambiguous stimuli in more or less the same way, whereas above-average readers linger over ambiguous stimuli while categorizing endpoint stimuli more rapidly.

Other measures of reading ability. As a final post-hoc test, we explored which specific components of reading skill might underlie the relationship with performance on speech categorization tasks. To do this, we performed Pearson correlation tests between our estimates of psychometric slope, lapse rate, or the principal component that combines slope and asymptote parameters, and several tests targeting specific aspects of reading ability. Those correlations are summarized in Fig. 5. The most striking pattern is that the majority of significant correlations between test scores and our model parameters correlate with results from the ABX blocks, but not the single-stimulus blocks. Moreover, it appears that correlations between test scores and lapse rate are clustered on the dynamic-cue /ba/~/da/ continuum, whereas correlations with psychometric slope are clustered on the static-cue /ʃa/~/sa/ continuum. Perhaps unsurprisingly, the correlations between test scores and the principal component combining slope and asymptote parameters look something like a union of the correlation patterns for slope and lapse rate. These patterns suggest a striking difference between the ABX paradigm and the single-stimulus paradigm, in how performance on such tasks relates to reading ability.

Pairwise correlations of psychometric function parameters with tests targeting phonological processing were similar to correlations with real word and pseudoword reading measures (Fig. 5). However, a mixed model predicting psychometric slope as a function of phonological awareness (CTOPP Phonological Awareness) showed that after controlling for nonverbal IQ, the main effect of phonological awareness only showed a weak relationship to psychometric slope ($\beta = 0.018$, $SE = 0.009$, $p = 0.054$). Unsurprisingly, phonological awareness, as well as two other measures of phonological processing (phonological memory and rapid symbol naming) were significantly correlated with WJ-BRS in our sample (Fig. S2).

We therefore wanted to test the frequently discussed hypothesis that phonological awareness *mediates* the relationship between speech perception and reading ability. In other words, we asked whether our data supported the theory that abnormal labeling of phonemes is associated with poor phonological awareness, which in turn drives reading difficulty.

Following the mediation analysis approach of Baron and Kenny[74], we tested this hypothesis using the average psychometric slope measured for each individual as a predictor, the WJ-BRS as an outcome measure, and the CTOPP phonological awareness score as a mediator. Using the R mediate package[75] with 2000 bootstrapped simulations, the estimated indirect coefficient describing the mediation effect was significant ($\beta = 3.85$, $SE = 2.11$, $p = 0.011$). This analysis indicates that phonological awareness is a partial mediator (36.4% mediation) of the relationship between slope and WJ-BRS. However, when nonverbal IQ was included as a covariate, the estimated indirect coefficient did not meet the criterion of significance ($\beta = 1.6984$, $SE = 1.485$, $p = 0.071$), corresponding to a 25.5% mediation in the relationship between slope and WJ-BRS. Model comparison provided strong evidence for the inclusion of nonverbal IQ as a covariate ($F(1,41) = 9.66$, $p = 0.003$). Therefore, we did not observe clear evidence that phonological awareness mediates the relationship between slope and reading ability once nonverbal IQ is controlled for, although this effect deserves further investigation with a larger sample.

On the other hand, conducting this same analysis with the CTOPP Rapid Symbol Naming measure as the mediator between psychometric slope and reading ability indicated a larger mediation effect: even with the inclusion of nonverbal IQ as a covariate, this measure of rapid naming provided a 52.2% mediation of the relationship between slope and WJ-BRS ($\beta = 3.476$, $SE = 1.745$, $p = 0.021$).

One interpretation of these analyses is that, contrary to what has been argued in the literature[59,76,77], the phoneme categorization task we used is not in fact a reliable probe of phonological awareness (whereas the various test batteries are). The phoneme categorization task may more strongly reflect the contributions of rapid access to phoneme labels, not awareness of phonemes themselves. If, on the other hand, one accepts that (psychometric
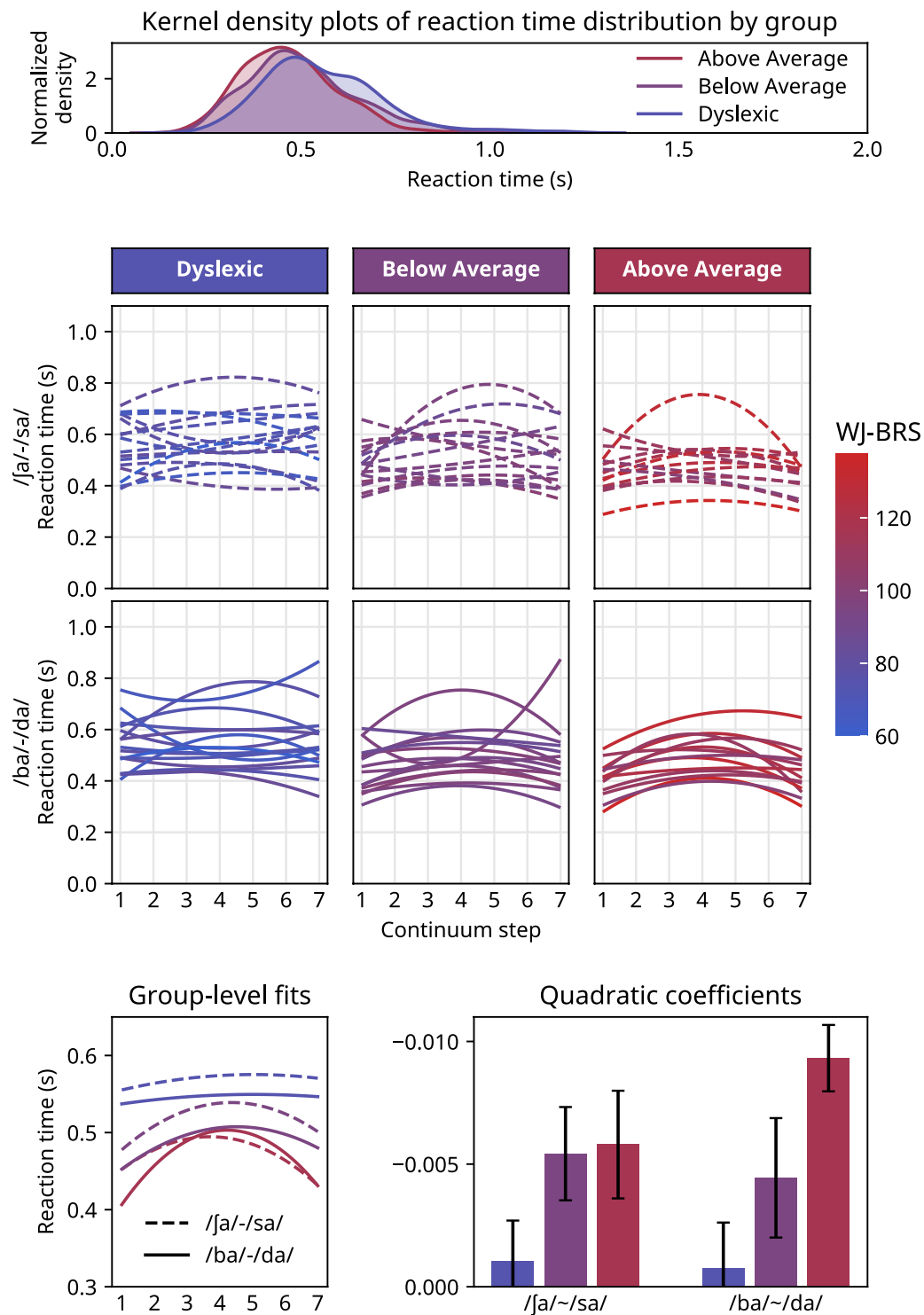
Figure 4. (**A**) Gaussian kernel density plots (75 ms bandwidth) of raw reaction times pooled across continua and paradigms for each group. (**B**) Polynomial fits to reaction times for both speech continua. (**C**) Polynomials fit to aggregated participant reaction times for each of the two speech continua. (**D**) Mean quadratic coefficients (±1 standard error of the mean) of the polynomial fits for subjects in each group for the two speech continua.

models of) children's performance on phoneme categorization tasks do reflect differences in phonological processing, then these results call into question the specificity of the component tests of these various diagnostic batteries.
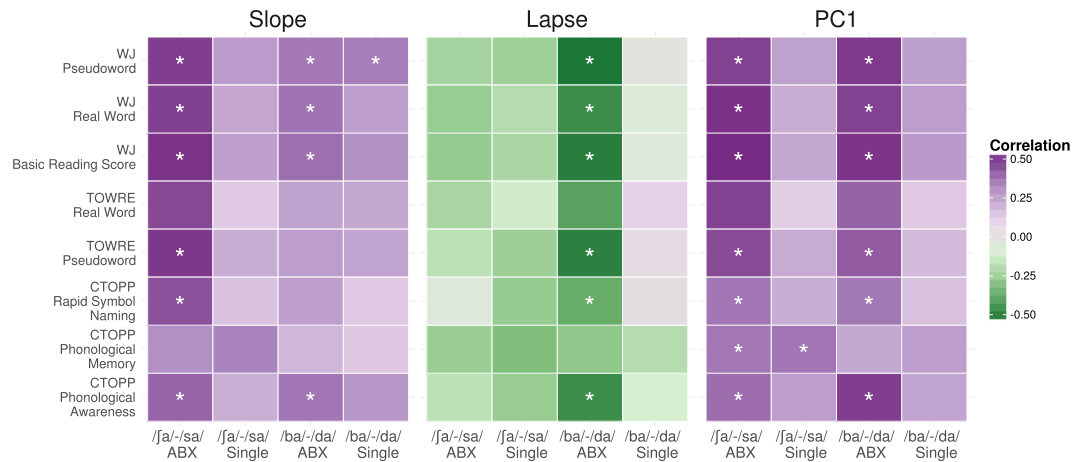
Figure 5. Pearson correlations between behavioral measures and slope, lapse, and first principal component (PC1) of the psychometric functions estimated from behavioral responses. Color gradient indicates the magnitude of the correlation and asterisks indicate significant correlations (p < 0.05, corrected for multiple tests using false discovery rate estimation).

Discussion

Many previous studies have claimed to show a relationship between dyslexia and poor categorical perception of speech phonemes (particularly for contrasts that rely on dynamic spectrotemporal cues), while others have suggested that the apparent auditory or linguistic processing impairments are in fact the result of general attention or memory deficits that manifest because of task difficulty. Our data unambiguously show a relationship between reading ability and phoneme categorization performance: higher reading scores were associated with steeper categorization functions, lower lapse rates, faster response times, and greater tendency to respond more quickly for unambiguous stimuli. But we also set out to answer two more specific questions: first, do poor readers struggle to categorize sounds on the basis of dynamic auditory cues to a greater extent than static cues? And second, are poor readers affected by task difficulty more than typically developing children? We also took care to show the ramifications of methodological choices when modeling categorization data as a psychometric function (particularly the common assumption of zero lapse rate). We discuss each of these points in turn.

It has been argued that dyslexics suffer primarily from an auditory processing deficit that specifically affects the processing of temporal information, leading to unique impairments in processing dynamic temporal and spectrotemporal cues such as formant transitions[43,55] or modulations on the timescale of 2 to 20 Hz[39]. We found that poor readers showed, on average, less steep categorization functions regardless of whether they were categorizing on the basis of static or dynamic speech cues compared to strong readers. This was true regardless of whether reading ability was treated as a continuous variable, or if the Dyslexic children were compared to Below Average or Above Average reading groups. However, estimated lapse rate was related to cue type in our data (though in the opposite direction than would be predicted from the literature: higher lapse rates were associated with the static cue continuum, not the dynamic one). These results run counter to the findings of Vandermosten and colleagues (2010, 2011), who showed a greater deficit in categorization of speech continua involving dynamic cues compared to static cues (although note that they did not report reaction times or lapse rates, only psychometric slopes). We believe the key difference arises from our choice to equalize the duration of the cues in all stimulus tokens. Vandermosten and colleagues used a 100 ms dynamic cue in their /ba/~/da/ continuum, and a 350 ms static cue in their /i/-/y/ continuum, which raised the possibility that integration of sensory signals over time, not the static or dynamic nature of the speech cue, was the source of their observed difference.

One interpretation of our results is that categorizing phonemes may be more difficult when judgments must be made on the basis of very brief cues (whether static or dynamic), and that children with dyslexia are especially susceptible to this difficulty. However, most evidence suggests that slowing down speech does not make it more discriminable for reading and language impaired children[78,79]. Another possibility is that dyslexics have some impairment specifically in processing syllable-length sounds, as has been recently suggested[8,39,80], which cannot be ruled out based on our findings. What we can conclude is that modulations on the order of 10 Hz (the rate of formant transition changes in our /ba/~/da/ continuum) do not pose a unique difficulty for poor readers, when compared to static speech cues of comparable duration.

In our data, estimates of psychometric slope and lapse rate were unrelated to whether the task was the more difficult ABX paradigm or the easier single-stimulus paradigm. Additionally, the covariates for ADHD diagnosis and nonverbal IQ were never shown to contribute to better model fits of the relationship between standardized reading ability and psychometric shape, a contribution we would have expected if task difficulty were a serious issue for our participants. This result, in conjunction with our findings that poor readers have less release from effort at the endpoints of the continua (measured by reaction time) and do not show signs of becoming distracted during the course of trial blocks any more so than the other participants, argue against interpreting the categorical perception deficit primarily as a manifestation of inattention or ADHD. Our results are therefore in agreement with another study that showed abnormal identification curves in dyslexic children both with and

without ADHD[60]. To the extent that task difficulty may contribute at all to performance differences, our results are more in line with the hypothesis that complex tasks burden working memory in poor readers rather than that they drive inattention.

Our PCA analysis of overall psychometric curve shape provides some evidence that poor readers were more affected by task difficulty than strong readers, and the correlations between various tests of reading ability and psychometric slope or lapse rate were nearly always greater in the ABX paradigm than in the single-stimulus paradigm (see Fig. 5). However, the relationship between reading ability and both slope and lapse were best modeled without considering test paradigm, indicating that the deficit we observe in poor readers is not primarily driven by task difficulty. It is possible that our ABX and single-stimulus conditions were simply not different enough in difficulty to reveal an effect of task difficulty in our statistical models, despite the apparent difference suggested by the post-hoc correlations with test battery scores. Pertinent to that possibility is a recent meta-analysis of categorical perception studies among children with dyslexia[35], which found that there is generally a larger effect size reported for discrimination tasks than identification tasks. Indeed, during initial piloting of this study, subjects also performed same/different discrimination task with the same stimulus continua, but performance was so uniformly poor that the discrimination task was dropped from the final experiment. In any case, our findings reinforce claims that in order to fully understand what stimulus-dependent difficulties are faced by poor readers, the stimuli should be tested in more than one context to separate the effects of stimulus properties from the effects of task demands[25].

Our results did not provide clear evidence for the hypothesis that phonological awareness mediates the relationship between reading ability and the ability to label speech sounds categorically. However, our findings are in line with several other recent results. In a 2016 study of 49 dyslexic children and 86 controls, Hakvoort *et al.* performed a similar mediation analysis to test whether phonological awareness mediated the relationship between a measure of categorical perception and reading ability[81]. They did not find evidence for this relationship. Like us, they found that rapid automatic naming partially mediated the relationship between speech and reading measures. Additionally, another study of categorical perception deficits in dyslexic children found no correlation between phonological awareness and performance on phoneme discrimination and identification tasks[82]. Also relevant is a study which showed that dyslexics were both worse at perceiving speech in noise and worse at labeling phonemes, but did not find evidence that the categorical labeling mediates the relationship between speech-in-noise impairments and reading ability[83]. In sum, our study adds another piece of evidence to the view that the classic model relating speech perception to phonological awareness to reading ability may be too simplistic, insufficiently specified, or possibly even incorrect.

One important point that our study clarifies is the influence of methodological choices on psychometric fitting results. Specifically, we compared 4-parameter models of participants' psychometric functions with traditional 2-parameter models that ignore lapse rate by fixing the asymptotes at zero and one. Our cross-validation of the asymptote prior distributions make it clear that a 2-parameter model is a worse fit to the data than any of the possible 4-parameter models we tested (see Materials & Methods for further details). More importantly, using a 2-parameter model in spite of its poor fit to the data leads to incorrect conclusions: 2-parameter models indicated that slopes were lower for the static cue continuum, whereas 4-parameter models indicated that slope was no different between the two continua, but lapse rate was higher on the static cue continuum. In either case, the main finding—a clear relationship between reading ability and psychometric slope—remains. This is, to a degree, reassuring about previous conclusions drawn from 2-parameter models. Many prior studies are still probably correct in finding that children with dyslexia tend to have shallower categorization function slopes on particular types of stimuli. However, by failing to model lapse rate those studies may have overestimated the effect size when comparing children with dyslexia against typical readers.

To further illustrate this point, we explored how the conclusions of our study would have changed if we had changed the assumptions of our psychometric fitting routine. Specifically, we varied the width of the Bayesian priors governing the two asymptotic parameters from 0 (effectively fixed) to 0.5 in steps of 0.05 and observed the effects on the correlation of reading ability and slope. We fit psychometric functions for each subject at each of the eleven prior widths, and at each prior width, we computed the correlations between reading score and slope or lapse rate on the ABX trials. Figure 6 illustrates the result; the apparent correlation between reading ability and psychometric slope is greatest when the priors on the asymptotes are strongest, and the correlation diminishes when the asymptotes are allowed to vary more freely. Thus, it is critical to determine the optimal parameterization of the psychometric function before interpreting the resulting parameters of the fit.

What is still not clear, and is not within the scope of this study, is the extent to which low-level sensory abilities drive the relationship between phoneme categorization performance and low reading ability. Gap detection thresholds (the gold standard for measuring temporal resolution) have been shown to be normal in dyslexics[12,17,84,85], so poor auditory temporal resolution is not likely to be the underlying cause for the correlation between reading ability and categorization performance. There is some evidence that backward masking is more severe in dyslexics[86], which could potentially mask the consonant cue preceding the vowel in /ba/~/da/ and /ʃa/~/sa/ stimuli, and could explain the difference between our findings and those of Vandermosten and colleagues (2010, 2011) regarding static-cue stimuli. However, Rosen and Manganari[87] showed that differences in backward masking did not affect categorical perception, so there is to date no clear relationship between any measured low-level deficit among poor readers and their ability to categorize speech sounds.

It has been long established that perceptual boundaries typically sharpen as a function of linguistic experience[88], and categorical labeling of speech sounds in pre-lingual children has been shown to predict later success in reading[89]. Thus, impairment on categorization tasks among children with dyslexia may represent the lower end of a distribution of literacy and categorical perception abilities that occurs in the general population, rather than a specific deficit that is unique to dyslexia[69]. Our inclusion of children with a large range of reading abilities, rather than just two groups of children widely separated in literacy, suggested that there is no clear and distinct
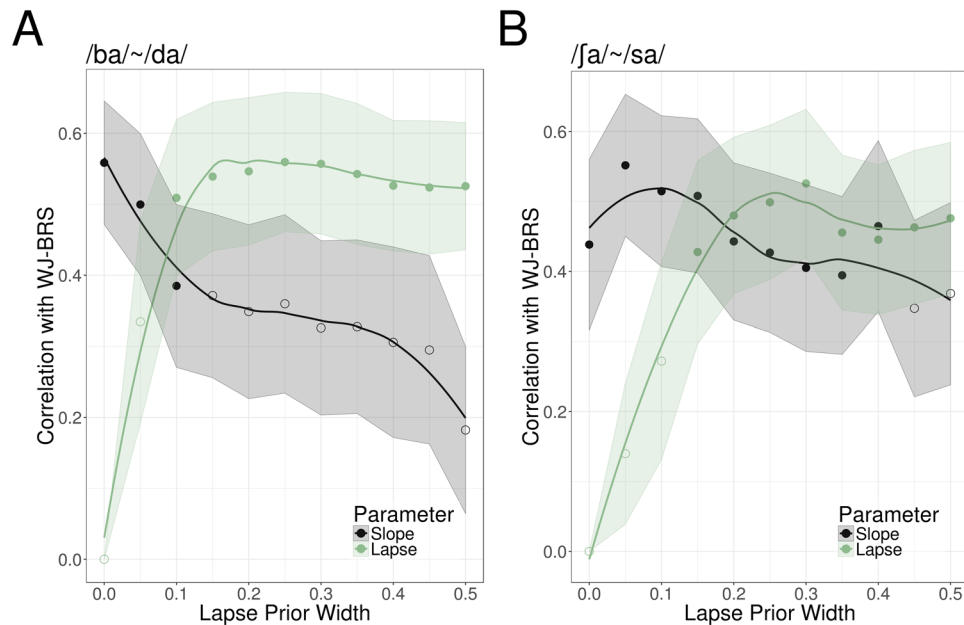
Figure 6. Correlations between slope, lapse rate, and reading ability (WJ-BRS score) as a function of the asymptotic prior width. Correlations are for psychometric functions fit to ABX trials. Error ribbons are computed by 10,000 bootstrap simulations. Filled dots indicate significant correlations ($p < 0.05$).

boundary between children with dyslexia and below-average readers that do not meet the criterion for reading disability. Our attempt to predict the group label for out-of-sample behavioral responses with QDA and LDA indicated that the groups were not easily separable by psychometric function shape. Although the mean slope and lapse parameters differed significantly by group, it was difficult to accurately predict the group an individual belonged to on the basis of their estimated slope or lapse. Moreover, variation in our models of children's perceptual boundaries was moderately related to reading ability, and the effect size and variability across our sample argues against interpreting the group-level effect as a unifying feature of dyslexia. Some poor readers had apparently normal psychometric functions, and there was considerable overlap in performance between poor readers in the Dyslexic group and merely Below Average readers. This aspect of our findings is in line with a multiple deficit model of dyslexia, whereby multiple underlying mechanisms, including auditory, visual and language deficits, all confer risk for reading difficulties[90–93].

Our results invite several questions for further exploration. First, although we found that children with dyslexia showed poor performance categorizing speech on the basis of (100 ms) steady-state spectral cues, Vandermosten et al.[43,55] showed that a large sample of dyslexic adults and children were not significantly impaired on a similar task with (350 ms) steady-state vowels. A natural extension of this work would be to parametrically vary cue duration, to look for a duration-dependent difference between readers with and without dyslexia. Such a difference could reflect a difference in sensory integration, which has been probed at the group level in dyslexics for visual stimuli, but not in relation to other sensory modalities[94]. Another question raised by our research is how directly auditory phoneme categorization relates to the common diagnostic measure "phonological awareness". Although it has been hypothesized that poor phoneme categorization reflects fuzzy underlying phonological representations, which in turn mediate reading ability, our study showed only weak correlations between categorization function slope and phonological awareness, compared to the relationship between slope and a compound measure of reading skill. However, our mediation analysis did suggest that measures of phonological awareness and/or rapid naming may partially mediate the relationship between phoneme categorization and reading skill. Therefore, it is worthwhile to perform a more targeted analysis of the relationship between categorical perception of speech, phonological awareness and reading ability.

Materials and Methods

Participants. A total of 52 native English-speaking school-aged children with normal hearing were recruited for the study in the study. Children ages 8–12, without histories of neurological or auditory disorders, were recruited from a database of volunteers in the Seattle area (University of Washington Reading & Dyslexia Research Database). Parents and/or legal guardians of all participants provided written informed consent under a protocol that was approved by the University of Washington Institutional Review Board. All experiments were performed in accordance with relevant guidelines and regulations of the Institutional Review Board. All subjects had normal or corrected-to-normal vision. Participants were tested on a battery of assessments, including the Woodcock-Johnson IV (WJ-IV) Letter Word Identification and Word Attack sub-tests, the Test of Word Reading Efficiency (TOWRE), and the Wechsler Abbreviated Scale of Intelligence (WASI-III). All participants underwent a hearing screening to ensure pure tone detection at 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz and 8000 Hz in both ears at 25 dB HL or better. A total of six subjects who were initially recruited did not pass the hearing screening and

	Dyslexic (15 subjects)	Below Average (16 subjects)	Above Average (13 subjects)	Significance (Dyslexic v. Below Average)	Significance (Dyslexic v. Above Average)
Males/Females	5m 10f	9m 7f	3m 10f		
Age	9.8 (1.4)	10.10 (1.0)	9.95 (1.4)	0.232	0.844
WASI-III					
FS-2	99.9 (14.3)	108.7 (12.4)	124.0 (15.4)	0.0293	0.0005*
Nonverbal IQ	46.0 (8.2)	54.1 (9.5)	59.61 (8.7)	0.008	0.0002*
WJ- IV					
WJ-BRS	74.6 (6.8)	93.8 (5.3)	117.2 (11.8)	2.184e-6*	5.584e-5*
Word reading	69.7 (10.5)	93.0 (6.8)	116.5 (12.6)	5.402e-6*	5.613e-6*
Nonword reading	84.2 (6.8)	97.5 (4.5)	115.2 (12.3)	1.446e-5*	6.780e-6*
TOWRE 2					
TOWRE Index	67.6 (7.10)	85.8 (9.87)	113.2 (9.06)	0.0002*	7.816e-6*
Word reading	67.4 (13.2)	88.5 (11.6)	115.3 (11.4)	0.0009*	8.459e-6*
Nonword reading	71.1 (3.2)	84.6 (9.1)	109.6 (11.1)	0.0003*	5.525e-6*
CTOPP 2					
Phonological awareness	86.6 (9.5)	95.8 (14.6)	107.23 (13.3)	0.109	0.0005*
Phonological memory	88.1 (17.5)	93.5 (10.1)	107.8 (17.5)	0.634	0.006
Rapid symbolic processing	75.0 (8.4)	82.6 (8.7)	98.6 (9.1)	0.068	1.108e-5*

Table 2. Summary statistics and group differences on various demographic, reading and cognitive measures. Summary statistics show means and (sample standard deviations). Significance assessed by the Wilcoxon test. Stars indicate p-values that meet threshold for significance (p < 0.05) after Bonferroni correction for 24 comparisons.

were not entered into the study, and two more were unable to successfully complete training (described below). Thus, a total of 44 children participated in the full experiment.

Demographics. In order to understand the relationship between phoneme categorization ability and reading ability, we selected our cohort of participants to encompass both impaired and highly skilled readers. Although we treat reading ability as a continuous covariate in our statistical analyses, for the purpose of recruitment and data visualization we defined three groups based on the composite Woodcock-Johnson Basic Reading Score (WJ-BRS) and TOWRE index. The "Dyslexic" group comprised participants who scored 1 standard deviation or more below the mean (standardized score of 100) on both the WJ-BRS and TOWRE index; Above Average readers were defined as those who scored above the mean on both tests; and Below Average readers were defined as participants who fell between the Dyslexia and Above Average groups. We used both the WJ-BRS and TOWRE index in our criterion to improve the confidence of our group assignments, though they are highly correlated measures in our sample (r = 0.89, p < 2e−16). There were 15 subjects in the Dyslexic group, 16 in the Below Average group, and 13 in the Above Average group. There were no significant differences in age between groups (Kruskal-Wallis rank sum test, H(2) = 1.0476, p = 0.5923), nor was there a significant correlation between age and WJ-BRS score (r = 0.02, p = 0.89). We did not exclude participants with ADHD diagnoses from the study because ADHD is highly comorbid with dyslexia[95,96], so this inclusion contributes to a more representative sample of children. However, we did account for the presence of ADHD diagnosis in our statistical models. Of our 44 participants, 7 had a formal diagnosis of ADHD: 2 in the Above Average group, 1 in the Below Average group, and 4 in the Dyslexic group.

Table 2 shows group comparisons on measures of reading and cognitive skills, and the distributions of each variable are illustrated in Supplementary Fig S1. Pairwise correlations between behavioral measures are shown in Fig. S2. All subjects had an IQ of at least 80 as measured by the WASI-III FS-2. Therefore, although there was a significant difference in IQ scores across groups, we were not concerned that abnormally low cognitive ability would prevent any child from performing the experimental task. While we expected IQ scores to be higher among stronger readers due to the verbal component of IQ assessment, we also noticed a significant difference in nonverbal IQ across the groups as measured on the WASI-III Matrix Reasoning test. To be certain our results were not confounded by this difference, we also included nonverbal IQ as a covariate in our statistical analyses to confirm the specificity of the relationship with reading skills as opposed to IQ.

Stimuli. Two 7-step speech continua were created using Praat v. 6.0.37[97], a /ba/~/da/ continuum and a /ʃa/~/sa/ continuum, chosen to probe categorization on the basis of dynamic and static auditory cues, respectively. In the /ba/~/da/ continuum, the starting frequency of the second vowel formant (F2) transition was varied. Individuals who are insensitive to this spectrotemporal modulation should have shallow psychometric functions on the /ba/~/da/ categorization task. The other continuum was /ʃa/~/sa/, in which the spectral shape of the fricative noise was varied between tokens (but was static for any given token). Both the /ʃa/~/sa/ fricative duration and the /ba/~/da/ F2 transition duration were 100 ms. Endpoint and steady-state formant values for the /ba/~/da/ continuum were measured from recorded syllables from an adult male speaker of American English. The spectral

peaks of the /ʃa/~/sa/ continuum were chosen from the spectral peaks measured in a recorded /sa/ spoken by the same talker. Spectrograms of the endpoints of both continua are shown in Supplemental Fig. S3.

The /ba/~/da/ continuum. Synthesis of the /ba/~/da/ continuum followed the procedure described in Winn and Litovsky[98]. Briefly, this procedure involves downsampling a naturally produced /ba/ token, extracting the first four formant contours via linear predictive coding, altering the F2 formant contour to change the starting frequency from 1085 Hz (/ba/) to 1460 Hz (/da/) in seven linearly-spaced steps, linearly interpolating F2 values for the ensuing 100 ms to make a smooth transition to the steady-state portion of the vowel (1225 Hz), and re-generating the speech waveform using Praat's source-filter synthesis (with a source signal extracted from a neutral vowel from the same talker). Because this procedure eliminates high-frequency energy in the signal at the downsampling step, the synthesized speech sounds are then low-pass filtered at 3500 Hz and combined with a version of the original /ba/ recording that had been high-pass filtered above 3500 Hz. This improves the naturalness of the synthesized sounds, while still ensuring that the only differences between continuum steps are in the F2 formant transitions.

The /ʃa/~/sa/ continuum. The /ʃa/~/sa/ continuum was created by splicing synthesized fricatives lasting 100 ms onto a natural /a/ token excised from a spoken /sa/ syllable. The duration of /a/ was scaled to 250 ms using Praat's implementaion of the PSOLA algorithm[99]. Synthesized fricatives contained three spectral peaks centered at 3000, 6000, and 8000 Hz. The bandwidths and amplitudes of the spectral peaks were linearly interpolated between continuum endpoints in seven steps, and the resulting spectra were used to filter white noise. To improve the naturalness of the synthesized fricatives, a gentle cosine ramp over 75 ms and fall over the final 20 ms was imposed on the fricative envelope. Aside from this onset/offset ramping (which was applied equally to all continuum steps), the contrastive cue (the amplitudes and bandwidths of the spectral peaks) was steady throughout the 100-ms duration of each fricative.

Procedure. Stimulus presentation and participant response collection was managed with PsychToolbox for MATLAB[100,101]. Auditory stimuli were presented at 75 dB SPL via circumaural headphones (Sennheiser HD 600). Children were trained to associate sounds from the two speech continua with animal cartoons on the left and right sides of the screen and to indicate their answers with right or left arrow keypresses. For the /ʃa/~/sa/ continuum, participants selected between pink and purple snakes. For the /ba/~/da/ continuum, participants selected between two different sheep cartoons. Throughout all blocks, each cartoon was always associated with the same stimulus endpoint. Two experimental conditions were presented: an ABX condition and a single-interval condition.

The ABX condition. In ABX blocks, participants performed a two-alternative forced-choice identification task in which they heard three stimuli with a 400 ms ISI and indicated whether the third stimulus (X) was more like the first (A) or second (B) stimulus. Stimuli A and B were always endpoints of the continuum, and during the endpoint presentations, the animals associated with those sounds lit up. Based on pilot data with adult participants, varying endpoint order trial-by-trial (i.e., interleaving /ba/-/da/-X and /da/-/ba/-X trials in the same block) was deemed very difficult, so for our young listeners we fixed the order of endpoint stimulus presentation within each block. The order of blocks was counterbalanced across participants. Each block contained 10 trials for each of the seven continuum steps, for a total of 70 trials per block.

The single interval condition. In single interval blocks, participants heard a single syllable and decided which category it belonged to by selecting an animal. No text labels were used, so participants learned to associate each animal with a continuum endpoint during a practice round. Two subjects (neither in the Dyslexia group) lost track of the animals associated with the endpoints after one successful practice round and were given brief additional instruction before successfully completing the task. Each block contained 10 trials for each step on the continuum, for a total of 70 trials.

There were thus six blocks total: an ABX /ba/~/da/ block, an ABX /ʃa/~/sa/ block, a second ABX /ba/~/da/ block, a second ABX /ʃa/~/sa/ block, a single-interval /ba/~/da/ block, and finally a single-interval /ʃa/~/sa/ block. Other than the counterbalancing of ABX endpoint order mentioned above, the order of blocks was the same for all participants. Practice rounds were administered before the first ABX round and before each of the single-interval blocks. In practice rounds, participants were asked to categorize only endpoint stimuli and were given feedback on every trial. Participants had to score at least 75% correct on the practice round to advance to the experiment and were allowed to repeat the practice blocks up to three times. Two children did not meet this criterion and were not included in the study (one from the Above Average group and one from the Dyslexic group).

Psychometric curve fitting. Modeling of response data was performed with Psignifit 4.0[67], a MATLAB toolbox that implements Bayesian inference to fit psychometric functions. We fit a logistic curve with four parameters, modeling the two asymptotes, the width of the psychometric, and the threshold. The width parameter was later transformed to the slope at the threshold value. Four-parameter fits of each experimental block were performed with each of 11 possible priors: a prior that fixed the lapse rate at zero (in line with the approach in most previous studies of auditory processing in dyslexia), and ten uniform distribution priors with lower bound zero and upper bound ranging from 5% to 50% in steps of 5% (e.g., a lapse rate of 20% corresponds to constraining the lower asymptote between 0 and 0.2, and the upper asymptote between 0.8 and 1). Next, the optimal prior was chosen using leave-one-out cross-validation. Specifically, for each of the 261 test blocks across all participants, psychometric curves were fit to 69 of the 70 data points in that block, and the likelihood of the participant's selection on the held-out data point was calculated under the model. This process was repeated for each of the 70 data points in each block, for each of the 11 possible prior widths. The estimated likelihoods of the held-out data
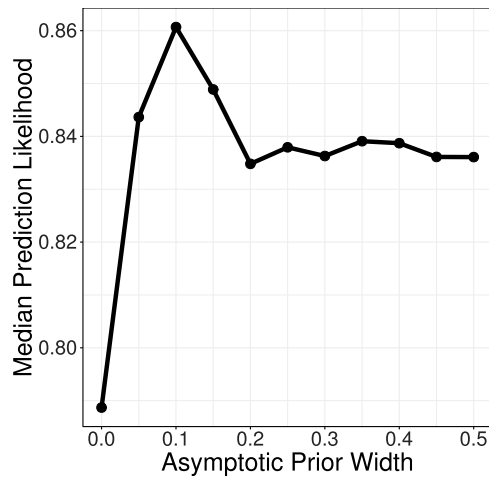
Figure 7. The median likelihood of held-out data points from all cross-validation trials is plotted against the width of the asymptotic parameter prior distributions (leave-one-out cross validation). The optimal prior width is clearly shown to be one that allows lapse rate to vary between 0 and 0.1. However, in all cases the 4-parameter model fit the data better than the 2-parameter model.

points were used as a goodness-of-fit metric and pooled across blocks and cross-validation runs to determine the median likelihood for each prior width. The optimal prior was determined to be a maximum lapse rate of 10%, as seen in Fig. 7; note that when the asymptotes were fixed at 0 and 1, the psychometric fits had the poorest fit to the data.

Having determined the optimal prior for the asymptote parameters, models were re-fit to obtain final estimates of the four parameters, for each combination of participant, cue continuum, and test paradigm. Additionally, 2-parameter fits (with asymptotes fixed at 0 and 1) were also recomputed, for later comparison with the optimal 4-parameter models. Any cases where the best-fit threshold parameter was not in the range of the stimulus continuum steps (1 through 7) were excluded from further analysis. Of the 261 total psychometric functions fit in the study, 12 of the 2-parameter fits and 10 of the 4-parameter fits were excluded on these grounds.

Reaction time analysis. Post-hoc analysis of reaction time data was also performed. To analyze reaction time, we first removed any trials with a reaction time shorter than 200 ms or longer than 2 s. The lower cutoff was based on the minimum response time to an auditory stimulus[102] such responses are assumed to be spurious or accidental button presses, and resulted in the exclusion of ~9% of trials. The upper cutoff was chosen based on the histograms of reaction times to exclude trials where subjects became distracted and were clearly temporarily disengaged from the experiment, and resulted in the exclusion of ~6% of trials. We fit each individual's remaining data as a function of continuum step with a 2nd order orthogonal polynomial basis (see Results for rationale), reading ability, and stimulus continuum with a mixed-effects regression model. For visualization, the reaction time data from each experimental block for each participant were reduced by fitting a χ^2 distribution to the reaction times at each continuum step, and using the peak of the fitted distribution as our summary statistic (because reaction time distributions tend to be right-skewed, making the mean a poor measure of central tendency).

Data Availability Statement

The de-identified datasets reported and analyzed in the current study, as well as the code for stimulus presentation, data analysis, and figure generation, is available at https://github.com/YeatmanLab/Speech_contrasts_public.

References

1. Lyon, G. R., Shaywitz, S. E. & Shaywitz, B. A. A definition of dyslexia. *Ann. Dyslexia* **53**, 1–14 (2003).
2. Shaywitz, S. E. Dyslexia. *N. Engl. J. Med.* **338**, 307–312 (1998).
3. Snowling, M. J. *Dyslexia, 2nd Edition. Wiley-Blackwell* (2000).
4. Snowling, M. J. Dyslexia as a Phonological Deficit: Evidence and Implications. *Child Psychol. Psychiatry Rev.* **3**, 4–11 (1998).
5. Tallal, P. Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang.* **9**, 182–198 (1980).
6. Farmer, M. E. & Klein, R. M. The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychon. Bull. Rev.* **2**, 460–493 (1995).
7. V Ingelghem, *et al.* An auditory temporal processing deficit in children with dyslexia. In *Learning disabilities: A challenge to teaching and instruction. Series: Studia Paedagogica* 47–63 (2005).
8. Goswami, U. A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* **15**, 3–10 (2011).
9. Steinbrink, C., Zimmer, K., Lachmann, T., Dirichs, M. & Kammer, T. Development of rapid temporal processing and its impact on literacy skills in primary school children. *Child Dev.* **85**, 1711–1726 (2014).
10. Poelmans, H. *et al.* Reduced sensitivity to slow-rate dynamic auditory information in children with dyslexia. *Res. Dev. Disabil.* **32**, 2810–2819 (2011).
11. Menell, P., McAnally, K. I. & Stein, J. F. Psychophysical sensitivity and physiological response to amplitude modulation in adult dyslexic listeners. *J. Speech, Lang. Hear. …* **42**, 797–803 (1999).
12. McAnally, K. I. & Stein, J. F. Scalp potentials evoked by amplitude-modulated tones in dyslexia. *J. Speech. Lang. Hear. Res.* **40**, 939–945 (1997).

13. Rocheron, I., Lorenzi, C., Füllgrabe, C. & Dumont, A. Temporal envelope perception in dyslexic children. *Neuroreport* **13**, 1683–7 (2002).
14. Witton, C. Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Curr. Biol.* **8**, 791–797 (1998).
15. Hämäläinen, J. A. *et al.* Common variance in amplitude envelope perception tasks and their impact on phoneme duration perception and reading and spelling in Finnish children with reading disabilities. *Appl. Psycholinguist.* **30**, 511–530 (2009).
16. Witton, C., Stein, J. F., Stoodley, C. J., Rosner, B. S. & Talcott, J. B. Separate influences of acoustic AM and FM sensitivity on the phonological decoding skills of impaired and normal readers. *J. Cogn. Neurosci.* **14**, 866–874 (2002).
17. Boets, B., Wouters, J., van Wieringen, A. & Ghesquière, P. Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: A longitudinal study of the auditory temporal processing theory. *Neuropsychologia* **45**, 1608–1620 (2007).
18. Stoodley, C. J., Hill, P. R., Stein, J. F. & Bishop, D. V. M. Auditory event-related potentials differ in dyslexics even when auditory psychophysical performance is normal. *Brain Res.* **1121**, 190–199 (2006).
19. Gibson, L. Y., Hogben, J. H. & Fletcher, J. Visual and auditory processing and component reading skills in developmental dyslexia. *Cogn. Neuropsychol.* **23**, 621–642 (2006).
20. Dawes, P. *et al.* Temporal auditory and visual motion processing of children diagnosed with auditory processing disorder and dyslexia. *Ear Hear.* **30**, 675–686 (2009).
21. Thomson, J. M. & Goswami, U. Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *J. Physiol. Paris* **102**, 120–129 (2008).
22. Goswami, U., Fosker, T., Huss, M., Mead, N. & SzuCs, D. Rise time and formant transition duration in the discrimination of speech sounds: The Ba-Wa distinction in developmental dyslexia. *Dev. Sci.* **14**, 34–43 (2011).
23. Thomson, J. M., Fryer, B., Maltby, J. & Goswami, U. Auditory and motor rhythm awareness in adults with dyslexia. *J. Res. Read.* **29**, 334–348 (2006).
24. Banai, K. & Ahissar, M. Poor frequency discrimination probes dyslexics with particularly impaired working memory. *Audiol. Neuro-Otology* **9**, 328–340 (2004).
25. Banai, K. & Ahissar, M. Auditory processing deficits in dyslexia: Task or stimulus related? *Cereb. Cortex* **16**, 1718–1728 (2006).
26. Bus, A. G. & van IJzendoorn, M. H. Phonological awareness and early reading: A meta-analysis of experimental training studies. *J. Educ. Psychol.* **91**, 403–414 (1999).
27. Hulme, C. *et al.* Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *J. Exp. Child Psychol.* **82**, 2–28 (2002).
28. Merzenich, M. M. *et al.* Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* **271**, 77–81 (1996).
29. Tallal, P. *et al.* Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science (80-.).* **271**, 81–84 (1996).
30. Rosen, S. Language disorders: A problem with auditory processing? *Curr. Biol.* **9** (1999).
31. Rosen, S. Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics* **31**, 509–527 (2003).
32. Ramus, F. *et al.* Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain* **126**, 841–865 (2003).
33. Serniclaes, W., Van Heghe, S., Mousty, P., Carré, R. & Sprenger-Charolles, L. Allophonic mode of speech perception in dyslexia. *J. Exp. Child Psychol.* **87**, 336–361 (2004).
34. Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S. & Sprenger-Charolles, L. Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *J. Exp. Child Psychol.* **101**, 137–155 (2008).
35. Noordenbos, M. W. & Serniclaes, W. The Categorical Perception Deficit in Dyslexia: A Meta-Analysis. *Sci. Stud. Read.* **19**, 340–359 (2015).
36. Calcus, A., Deltenre, P., Colin, C. & Kolinsky, R. Peripheral and central contribution to the difficulty of speech in noise perception in dyslexic children. *Dev. Sci.* **21** (2018).
37. Dole, M., Hoen, M. & Meunier, F. Speech-in-noise perception deficit in adults with dyslexia: Effects of background type and listening configuration. *Neuropsychologia* **50**, 1543–1552 (2012).
38. Ziegler, J. C., Pech-Georgel, C., George, F. & Lorenzi, C. Speech-perception-in-noise deficits in dyslexia. *Dev. Sci.* **12**, 732–745 (2009).
39. Law, J. M., Vandermosten, M., Ghesquière, P. & Wouters, J. The relationship of phonological ability, speech perception, and auditory perception in adults with dyslexia. *Front. Hum. Neurosci.* **8** (2014).
40. Boets, B. *et al.* Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Res. Dev. Disabil.* **32**, 560–570 (2011).
41. Hämäläinen, J. A., Salminen, H. K. & Leppänen, P. H. T. Basic Auditory Processing Deficits in Dyslexia. *J. Learn. Disabil.* **46**, 413–427 (2013).
42. Goswami, U. *et al.* Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proc. Natl. Acad. Sci.* **99**, 10911–10916 (2002).
43. Vandermosten, M. *et al.* Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proc. Natl. Acad. Sci.* **107**, 10389–10394 (2010).
44. Siegel, L. S. & Ryan, E. B. Subtypes of developmental dyslexia: The influence of definitional variables. *Read. Writ.* **1**, 257–287 (1989).
45. Swanson, H. L. Working memory in learning disability subgroups. *J. Exp. Child Psychol.* **56**, 87–114 (1993).
46. Vargo, F. E., Grosser, G. S. & Spafford, C. S. Digit span and other WISC-R scores in the diagnosis of dyslexia in children. *Percept. Mot. Skills* **80**, 1219–29 (1995).
47. Wang, S. & Gathercole, S. E. Working memory deficits in children with reading difficulties: Memory span and dual task coordination. *J. Exp. Child Psychol.* **115**, 188–197 (2013).
48. Amitay, S., Ben-Yehudah, G., Banai, K. & Ahissar, M. Disabled readers suffer from visual and auditory impairments but not from a specific magnocellular deficit. *Brain* **125**, 2272–2285 (2002).
49. Ramus, F. & Szenkovits, G. What phonological deficit? In *Quarterly Journal of Experimental Psychology* **61**, 129–141 (2008).
50. Ahissar, M., Lubin, Y., Putter-Katz, H. & Banai, K. Dyslexia and the failure to form a perceptual anchor. *Nat. Neurosci.* **9**, 1558–1564 (2006).
51. Ahissar, M. Dyslexia and the anchoring-deficit hypothesis. *Trends Cogn. Sci.* **11**, 458–465 (2007).
52. Ziegler, J. C. Better to lose the anchor than the whole ship. *Trends in Cognitive Sciences* **12**, 244–245 (2008).
53. Boets, B. *et al.* Intact but less accessible phonetic representations in adults with dyslexia. *Science (80-.).* **342**, 1251–1254 (2013).
54. Roach, N. W., Edwards, V. T. & Hogben, J. H. The tale is in the tail: An alternative hypothesis for psychophysical performance variability in dyslexia. *Perception* **33**, 817–830 (2004).
55. Vandermosten, M. *et al.* Impairments in speech and nonspeech sound categorization in children with dyslexia are driven by temporal processing difficulties. *Res. Dev. Disabil.* **32**, 593–603 (2011).
56. Wichmann, F. A. & Hill, N. J. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* **63**, 1293–1313 (2001).

57. Wichmann, F. A. & Hill, N. J. The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept. Psychophys.* **63**, 1314–1329 (2001).
58. Reed, M. A. Speech perception and the discrimination of brief auditory cues in reading disabled children. *J. Exp. Child Psychol.* **48**, 270–292 (1989).
59. Manis, F. R. *et al.* Are Speech Perception Deficits Associated with Developmental Dyslexia? *J. Exp. Child Psychol.* **66**, 211–235 (1997).
60. Breier, J. I. *et al.* Perception of Voice and Tone Onset Time Continua in Children with Dyslexia with and without Attention Deficit/Hyperactivity Disorder. *J. Exp. Child Psychol.* **80**, 245–270 (2001).
61. Chiappe, P., Chiappe, D. L. & Siegel, L. S. Speech Perception, Lexicality, and Reading Skill. *J. Exp. Child Psychol.* **80**, 58–74 (2001).
62. Maassen, B., Groenen, P., Crul, T., Assman-Hulsmans, C. & Gabreëls, F. Identification and discrimination of voicing and place-of-articulation in developmental dyslexia. *Clin. Linguist. Phonetics* **15**, 319–339 (2001).
63. Zhang, Y. *et al.* Universality of categorical perception deficit in developmental dyslexia: An investigation of Mandarin Chinese tones. *J. Child Psychol. Psychiatry Allied Discip.* **53**, 874–882 (2012).
64. Messaoud-Galusi, S., Hazan, V. & Rosen, S. Investigating Speech Perception in Children With Dyslexia: Is There Evidence of a Consistent Deficit in Individuals? *J. Speech Lang. Hear. Res.* **54**, 1682 (2011).
65. Noordenbos, M. W., Segers, E., Serniclaes, W. & Verhoeven, L. Neural evidence of the allophonic mode of speech perception in adults with dyslexia. *Clin. Neurophysiol.* **124**, 1151–1162 (2013).
66. Collet, G. *et al.* Effect of phonological training in French children with SLI: Perspectives on voicing identification, discrimination and categorical perception. *Res. Dev. Disabil.* **33**, 1805–1818 (2012).
67. Schütt, H., Harmeling, S., Macke, J. H. & Wichmann, F. A. Psignifit 4: Pain-free Bayesian Inference for Psychometric Functions. *J. Vis.* **15**, 474 (2015).
68. Treutwein, B. & Strasburger, H. Fitting the psychometric function. *Percept. Psychophys.* **61**, 87–106 (1999).
69. Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M. & Makuch, R. Evidence That Dyslexia May Represent the Lower Tail of a Normal Distribution of Reading Ability. *N. Engl. J. Med.* **326**, 145–150 (1992).
70. Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
71. Hastie, T., Tibshirani, R. & Friedman, J. The Elements of Statistical Learning. *Elements* **1**, 337–387 (2009).
72. Pisoni, D. B. & Tash, J. Reaction times to comparisons within and across phonetic categories. *Percept. Psychophys.* **15**, 285–290 (1974).
73. Repp, B. H. Perceptual equivalence of two kinds of ambiguous speech stimuli. *Bull. Psychon. Soc.* **18**, 12–14 (1981).
74. Baron, R. & Kenny, D. The moderator-mediator variable distinction in social psychological research. *J. Pers. Soc. Psychol.* **51**, 1173–1182 (1986).
75. Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. Package 'mediation'. *CRAN*, https://doi.org/10.1037/a0020761 (2015).
76. Joanisse, M. F., Manis, F. R., Keating, P. & Seidenberg, M. S. Language Deficits in Dyslexic Children: Speech Perception, Phonology, and Morphology. *J. Exp. Child Psychol.* **77**, 30–60 (2000).
77. Boets, B. *et al.* Towards a further characterization of phonological and literacy problems in Dutch-speaking children with dyslexia. *Br. J. Dev. Psychol.* **28**, 5–31 (2010).
78. Bradlow, A. R. *et al.* Effects of lengthened formant transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children. *J. Acoust. Soc. Am.* **106**, 2086 (1999).
79. Stollman, M. H. P., Kapteyn, T. S. & Sleeswijk, B. W. Effect of time-scale modification of speech on the speech recognition threshold in noise for hearing-impaired and language-impaired children. *Scand. Audiol.* **23**, 39–46 (1994).
80. Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D. & Giraud, A. L. Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron* **72**, 1080–1090 (2011).
81. Hakvoort, B. *et al.* The role of categorical speech perception and phonological processing in familial risk children with and without dyslexia. *J. Speech, Lang. Hear. Res.* **59** (2016).
82. Robertson, E. K., Joanisse, M. F., Desroches, A. S. & Ng, S. Categorical speech perception deficits distinguish language and reading impairments in children. *Dev. Sci.* **12**, 753–767 (2009).
83. Calcus, A., Lorenzi, C., Collet, G., Colin, C. & Kolinsky, R. Is There a Relationship Between Speech Identification in Noise and Categorical Perception in Children With Dyslexia? *J. Speech Lang. Hear. Res.* **59**, 835 (2016).
84. Adlard, A. & Hazan, V. Speech Perception in Children With Specific Reading Difficulties (Dyslexia). *Q. J. Exp. Psychol. Sect. A* **51**, 153–177 (1998).
85. Schulte-Körne, G., Deimel, W., Bartling, J. & Remschmidt, H. Role of auditory temporal processing for reading and spelling disability. *Percept. Mot. Skills* **86**, 1043–1047 (1998).
86. Wright, B. A. *et al.* Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature* **387**, 176–178 (1997).
87. Rosen, S. & Mangari, E. Speech and Nonspeech Auditory Processing in Children With Dyslexia? *J. speech, Lang. Hear. Res.* **44**, 720–736 (2001).
88. Garnica, O. K. The development of phonemic speech perception. in *Cognitive development and the acquisition of language*, 215–222 (1973).
89. Bradley, L. & Bryant, P. E. Categorizing sounds and learning to read - A causal connection. *Nature* **301**, 419–421 (1983).
90. Peterson, R. L. & Pennington, B. F. Developmental Dyslexia. *Annu. Rev. Clin. Psychol.* **11**, 283–307 (2015).
91. Pennington, B. F. From single to multiple deficit models of developmental disorders. *Cognition* **101**, 385–413 (2006).
92. Joo, S. J., Donnelly, P. M. & Yeatman, J. D. The causal relationship between dyslexia and motion perception reconsidered. *Sci. Rep.* **7** (2017).
93. Joo, S. J., White, A. L., Strodtman, D. J. & Yeatman, J. D. Optimizing text for an individual's visual system: The contribution of crowding to reading difficulties. *Cortex*, 1–27, https://doi.org/10.1016/j.cortex.2018.03.013 (2017).
94. Talcott, J. B., Hansen, P. C., Assoku, E. L. & Stein, J. F. Visual motion sensitivity in dyslexia: Evidence for temporal and energy integration deficits. *Neuropsychologia* **38**, 935–943 (2000).
95. Light, J. G., Pennington, B. F., Gilger, J. W. & DeFries, J. C. Reading disability and hyperactivity disorder: Evidence for a common genetic etiology. *Dev. Neuropsychol.* **11**, 323–335 (1995).
96. Stevenson, J. *et al.* Attention deficit hyperactivity disorder with reading disabilities: Preliminary genetic findings on the involvement of the ADRA2A gene. *J. Child Psychol. Psychiatry Allied Discip.* **46**, 1081–1088 (2005).
97. Boersma, P. Praat, a system for doing phonetics by computer. *Glot Int.* **5**, 341–347 (2001).
98. Winn, M. B. & Litovsky, R. Y. Using speech sounds to test functional spectral resolution in listeners with cochlear implants. *J. Acoust. Soc. Am.* **137**, 1430–1442 (2015).
99. Moulines, É. & Charpentier, F. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Commun.* **9**, 453–467 (1990).
100. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
101. The Mathworks Inc. MATLAB – MathWorks, 2016-11-26, www.mathworks.com/products/matlab (2016).
102. Fry, D. B. Simple Reaction-Times to Speech and Non-Speech Stimuli. *Cortex* **11**, 355–360 (1975).

Acknowledgements

Author Contributions

G.E.O., D.R.M. and J.D.Y. designed the experiments. G.E.O. and E.C.K. collected the data. G.E.O., D.R.M. and J.D.Y. wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-34823-8.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.