JSLHR Papers in Press. Published July 27, 2009, as doi: 10.1044/1092-4388(2009/08-0220)
The final version is at http://jslhr.asha.org.

Title: Speech perception abilities of adults with dyslexia: is there any evidence for a true deficit?

Authors: Valerie Hazan[1], Souhila Messaoud-Galusi[1], Stuart Rosen[1], Suzan Nouwens[2] and Bethanie Shakespeare[1]

Institutional affiliation:

[1] Speech Hearing and Phonetic Sciences, UCL, London, UK

[2] Institute of Linguistics OTS, University of Utrecht, Utrecht, the Netherlands

Corresponding Author:

Prof. Valerie Hazan, UCL Department of Speech Hearing and Phonetic Sciences, UCL, Chandler House, 2, Wakefield Street, London London WC1N 1PF
Tel: + 44 20 7679 4076 , Fax: + 44 (0) 207 679 4010

Running head: Speech perception in adults with dyslexia

Abstract

Purpose: This study investigated whether adults with dyslexia show evidence of a consistent speech perception deficit by testing phoneme categorization and word perception in noise.

Method: Seventeen adults with dyslexia and 20 average readers underwent a test battery including standardized reading, language and phonological awareness tests, and tests of speech perception. Categorization of a 'pea'/'bee' voicing contrast was evaluated using adaptive identification and discrimination tasks, presented in quiet and in noise, and a fixed-step discrimination task. Two further tests of word perception in noise were presented.

Results: There were no significant group differences for categorization in quiet or noise, for across- and within-category discrimination as measured adaptively, or word perception, but average-readers showed better across- and within-category discrimination in the fixed-step discrimination. Individuals did not show consistent poor performance across related tasks.

Conclusions: The small number of group differences, and lack of consistent poor individual performance, suggests weak support for a speech perception deficit in dyslexia. It seems likely that at least some poor performances are attributable to non-sensory factors like attention. It may also be that some individuals with dyslexia have speech perceptual acuity that is at the lower end of the normal range and exacerbated by non-sensory factors.

Introduction

Developmental dyslexia is a specific learning disability that is characterized by difficulties in reading and writing despite adequate intelligence, cognitive abilities and learning environments (Shaywitz et al., 1998; Snowling, 2000). Over the last thirty years, deficits in many aspects of auditory, speech perceptual and phonological processing have been identified in children and adults with dyslexia (for a review, see Ramus, Rosen, Dakin, Day, Castellote, White, and Frith, 2003). Here, we specifically address claims of a speech perceptual deficit in adults with dyslexia using a range of tests that tap individuals' ability to identify and discriminate minimal phonetic contrasts and their perception of speech in noise.

Developmental dyslexia is a deficit that continues to affect individuals in adulthood, and investigating the speech and language processing abilities of adults with dyslexia can be particularly informative as lapses in attention, which can affect performance on repetitive perceptual tasks in children (Davis, Castles, McAnally, & Gray, 2001; Moore, Ferguson, Halliday and Riley, 2008), are likely to be less prevalent in adults. In both adults and children, evidence of poor performance on phonological awareness tasks is rather pervasive (e.g., Elbro, Nielsen & Petersen, 1994, Snowling, Nation, Moxham, Gallagher & Frith, 1997; Ramus et al., 2003; Ziegler and Goswami, 2005; Snowling, 2000, although see also Reid, Szczerbinski, Iskierka-Kasperek & Hansen, 2007 for cases of individuals with dyslexia who have unimpaired phonological awareness). However, there is increasing debate as to whether poor performance on phonological awareness tasks reflects impoverished phonological representations or rather difficulties with the access or manipulation of these representations. For example, Szenkovitz and Ramus (2008) found that French adults with dyslexia performed well on tasks such as voicing assimilation, that require underlying phonological processes, even though they performed poorly at phonological tasks such as nonword repetition or phoneme deletion. They argue that phonological representations in individuals with dyslexia are in fact intact and that it is the access to these representations which is impaired, with poor performance exacerbated in tasks that impose a heavy short-term memory load.

If individuals with dyslexia do have impoverished phonological representations, then it would be expected that they should show deficits in tasks that require them to consistently assign speech sounds to phonemic categories, or that require them to determine whether acoustically-similar speech sounds belong to the same category. Early studies of phonemic categorization in adults and children with dyslexia were heavily influenced by the work of Tallal which suggested that children with dyslexia had particular difficulty with rapid temporal processing (Tallal, 1980). These early studies typically focused on the perception of synthesized phonemic contrasts that were cued by rapid formant transitions (e.g. ba/da contrasts), and presented these in identification and discrimination tasks. Studies with dyslexic adults generally found systematic small differences in phonetic perception, with the slopes of the identification function, a straightforward index of consistency in labeling, typically shallower in the dyslexic group (e.g., Steffens, Eilers, Grossglenn & Jallad**,** 1992). In discrimination tasks, Steffens et al. argued that adults with dyslexia lacked the 'degree of precision' shown by average readers in controlled laboratory tests. More recent studies on the categorization of phonemic contrasts in adults with developmental dyslexia have tended to confirm this pattern of a lower degree of consistency in phoneme identification, which results in shallower slopes of the identification function (Schwippert and Koopmans-van Beinum, 1998; van Beinum, Schwippert, Been, van Leeuwen, Kuijpers, 2005). Many studies of speech perception in children with dyslexia mirror this finding (e.g. Godfrey, Syrdal-Lasky, Millay & Knox, 1981; Reed, 1989; Manis et al., 1997; Werker and Tees, 1987; Breier et al, 2001; Boada and Pennington, 2006). However, some studies have failed to find significant group differences in identification between individuals with dyslexia and average readers in studies with children (e.g., Mody, Studdert-Kennedy & Brady, 1997**;** Adlard and Hazan, 1998; Joanisse, Manis, Keating & Seidenberg, 2000; Maassen, Groenen, Crul, Assman-Hulsmans & Gabreels, 2001; Blomert, Mitterer and Paffen, 2004; Robertson, Joanisse, Desroche and Ng, in press) and adults (Ramus et al., 2003).

Generally, in the dyslexia literature, it is increasingly recognized that it is not sufficient to show that significant group differences occur at some level of processing, but that it must also be shown that a substantial number of individuals show a performance that differs significantly from the norm (Ramus et al., 2003; Reid et al., 2007; Heath, Bishop, Hogben & Roach, 2006; Ziegler, Castel, Pech-Georgel, George, Alario & Perry, 2008; McArthur and Hogben, 2001; McArthur, Ellis, Atkinson & Coltheart, 2008). In studies with dyslexic children and adults that have reported individual data, there is ample evidence of significant individual differences in performance on speech perception tasks. For example, Adlard and Hazan (1998) found that only about a third of the 13 children with dyslexia that they tested showed evidence of consistent 'perceptual weakness' across different perceptual tasks, while the rest performed within norms on a majority or all of the tasks. Lieberman et al. (1985) found high error rates on a consonant perception task for 28% of their adults with dyslexia, with 22% performing within norms. Evidence of clear individual differences in adults with dyslexia was also reported by Steffens et al. (1992). Ramus et al. (2003) tested dyslexic adults on an extensive range of tasks tapping their phonological, auditory, visual and speech perceptual abilities, and aggregated performance on related tasks to obtain scores for 'rapid' and 'slow' auditory/speech processing and 'speech' scores to compare with 'nonspeech' scores. They found no evidence of significant group effects for tasks involving rapid auditory processing, and the dyslexic group did not perform significantly worse on speech than non-speech tasks. However, their scrutiny of individual results showed that 7 out of the 16 dyslexic participants (44%) and one out of 16 controls (6%) showed deviant performance on the ‘rapid’ aggregate scores and 6 dyslexic participants (37%) and one control (6%) on the ‘slow’ auditory/speech tasks. Five participants in the dyslexic group (31%) showed deviant performance in the ‘nonspeech’ task as opposed to 7 dyslexic participants (44%) and 2 controls (12%) in the ‘speech’ tasks. There was clearly heterogeneity within the dyslexic group and it should also be noted that some non-dyslexic individuals performed poorly on these experimental tasks, even if in lower proportion than in the dyslexic group. Ramus et al. (2003) concluded from their study that the cause of dyslexia is a phonological deficit, and that it may be accompanied in some individuals by additional visual, auditory or motor deficits.

If poorer performance on categorization tasks in at least some individuals with dyslexia does reflect poorly-specified phoneme representations, then it would follow that further degradations of the speech signal, such as that resulting from the addition of noise, should have a particularly deleterious effect on speech perception for these listeners (Ramus, 2001). Cornelissen, Hansen, Bradley & Stein (1996) investigated this hypothesis with dyslexic adults using a range of naturally-produced nonsense syllables covering a range of phoneme contrasts presented in different levels of white noise. They found similar patterns of consonant confusions across groups, with more sha/cha confusions made by the dyslexic group than controls. This pattern of poorer identification of CV items in noise in adults with dyslexia was also replicated in Ramirez and Mann (2005). A recent study of speech perception in noise also found a deficit in a group of children with dyslexia relative to reading and age controls using a range of naturally-produced nonsense syllables presented in different noise conditions (Ziegler, Pech-Georgel, George & Lorenzi, in press). However, the dyslexic group also showed normal masking release effects (i.e., better performance in fluctuating than in stationary noise) which led the authors to suggest that the poor performance in noise could not be attributed to poor temporal or frequency resolution, or to deficits in peripheral processing but rather that children with dyslexia are deficient in the ‘simultaneous integration of various speech cues required for robust speech identification’.

One alternative explanation of the heterogeneity seen in studies of the speech perception abilities of children and adults with dyslexia is that it is does not reflect a specific deficit in auditory or perceptual abilities but rather ‘errant task performance’, as caused by lapses in attention or confusion about the task procedure (Roach, Edwards & Hogben, 2004; Heath, Bishop, Hogben & Roach, 2006). Simulations of performance on adaptive discrimination tasks and categorical tasks that included errant trials yield patterns of group results that concur with those seen in studies of perceptual abilities in individuals with dyslexia (Roach et al., 2004; Davis, Castles, McAnally & Gray, 2001) although Breier et al. (2001) found a deficit in phoneme categorization in a group of dyslexic children whether or not they were diagnosed with ADHD. Roach et al. (2004) argue that in order to distinguish poor performance that is due to nonsensory factors from poor performance linked to a specific perceptual deficit, it is necessary to determine whether a task has construct validity, i.e. that it is tapping the dimension that is being investigated. This can best be done by showing that individuals that perform below norm on a particular task are also within the lower tail of the normal distribution for another task tapping that same dimension (Heath et al., 2006). Information about 'robustness' of poor performance on a specific task could also be gleaned by repeating the same task more than once with a given individual (Skottun and Skoyles, 2007), but this is very rarely done due to learning effects and to the use of already extensive test batteries in studies of dyslexia. Correlations across tasks tapping a similar perceptual ability have been examined in some studies of adults with dyslexia investigating auditory processing abilities (e.g., Talcott et al., 1998; Witton et al., 1998) but evidence for such construct validity for speech perceptual tasks is much scarcer (Ramus et al, 2003). However, as suggested by Heath et al. (2006), significant correlations across tasks should be interpreted with caution. Indeed, they argue that failure to find correlations between tasks can arise because some of the tasks are psychometrically weak. On the other hand, significant correlations may arise that are linked to task-related skills and abilities. Correlations are therefore more impressive if found across tasks that use different formats for assessing a given perceptual ability. However, even there, significant correlations do not imply that all individuals in the group are showing a consistent pattern of performance across tasks (e.g., Heath and Hogben, 2004). When considering whether individuals with a specific reading impairment have a perceptual deficit therefore, especially given evidence of within-group heterogeneity, the most reliable approach is to look at evidence of consistent poor performance for related tasks within individual participants rather than at group correlations. We argue that it is not necessarily the case that a listener who does not have a perceptual deficit will perform well in all related tasks, as all participants may show lapses in attention related to boredom or fatigue, especially in lengthy sessions involving a number of repetitive tasks. However, a participant who has a perceptual deficit should never be able to show within-norm performance on a test which is tapping the perceptual process that is deficient.

The aims of this study were therefore twofold. First, in order to investigate whether the poor performance of adults with dyslexia are due to specific perceptual deficits, participants were tested on both adaptive and fixed discrimination tasks tapping the same perceptual process. If poor discrimination is due to a specific perceptual deficit, we would expect performance in specific individuals to be consistently poor across these testing procedures for a given speech continuum. If it is linked to issues such as task difficulty or memory load, we might expect better performance in adaptive tasks that track a consistent level of accuracy for each individual than in fixed-step discrimination tasks which typically include a majority of presentations that are difficult to discriminate. Second, we hypothesize that if poor performance on identification or discrimination tasks does truly reflect the fact that adults with dyslexia have poorly-defined phonological representations, then performance on these tasks should be severely affected by the addition of noise, at least for individuals showing poor categorization abilities. A milder prediction is that if an individual has difficulty with a test in quiet, poor performance should be exacerbated in noise. To test this hypothesis, identification and discrimination tests for a /pi/-/bi/ (‘pea’/’bee’) voicing contrast were carried out both in quiet and in noise, and two additional tests of word perception in noise were also presented.

Method

Participants

Thirty-seven monolingual English native speakers aged between 18.02 and 31.11 years participated in the study. The adults, who were paid for their participation, were recruited through adverts to the student body at UCL and by contacting several dyslexia centres in London. Participants included 17 adults (10 men and 7 women) with a mean age of 22;10 years (s.d. 3;6 yrs) who had been diagnosed with dyslexia by a qualified educational psychologist at university or during their school years (DYS group). The average-reader (AR) group included 20 adults (8 men and 12 women) with a mean age of 23;5 years (s.d. 2.9 yrs) who had normal attainment in reading.

Adults who agreed to participate in the study were included if they passed a hearing screening (thresholds of 20 dB HL or better at 500, 1000, 2000 and 4000 Hz) if they were free of other developmental disorders (SLI, ADHD, autism, dyspraxia). They were required to score within a standard deviation of the standardized mean for TROG-2, a test of receptive grammar (Bishop, 2003). This criterion was used to exclude participants who might have had a language disability other than dyslexia or another language disability combined with dyslexia. All participants also had to score above -1 standard deviation of the standardized mean for verbal IQ (BPVS; Dunn, Dunn, Whetton, and Burley, 1997) and non-verbal IQ (WAIS- bloc design sub-section, Wechsler, 1997).

The participants’ reading level was assessed using the word and pseudoword reading lists of the TOWRE – Form A (Torgesen, Wagner, & Rashotte, 1999). Participants were instructed to read each list as fast as they could. The number of items read in 45 seconds provides a raw score. A standard score is then derived for the word and pseudoword reading lists, and a combined standard score is computed. All average readers scored above 90 and dyslexic readers below 90 on the standardized aggregate score of the TOWRE reading test. Mean data for these standardized tests are presented in Table 1.

[Table 1 about here]

The DYS and AR groups did not differ in terms of their age, non-verbal IQ and performance on a test of receptive grammar. As expected, the two groups differed on the word and pseudoword subtests, and aggregate score for the Test of Word Reading (TOWRE), indicating that the dyslexic group had a significantly lower reading level. The two groups also differed in terms of their verbal IQ.

Procedure

Ethical approval was granted by the UCL Research Ethics Committee. Participants were tested in a sound-treated room. Instructions and testing material were recorded by female native speakers of British English and were presented to participants through Sennheiser HD25-1 headphones. The sound level at which the stimuli were presented on the laptop computer was fixed for all listeners and identical to that used in our study with children (Messaoud-Galusi, Hazan and Rosen, 2007). The experiment took place over two sessions, each of an hour, with, for the majority of the participants, a few days in between sessions. Due to time constraints, a small number of participants had to complete the whole test battery in a single session.

Test battery

Standardized tests

Phonological Awareness: Phonological awareness was assessed using the rhyme and the spoonerism subtests of the Phonological Assessment Battery (PhAB) (Fredrickson, Frith & Reason, 1997). In the rhyme task, three words are presented orally and participants are required to repeat the two words that sound the same at the end (e.g., “sail, boot, nail” gives “sail, nail”). The first three trials are practice items for which feedback is given, followed by 21 test trials. The total number of correct responses is summed to obtain the final score.

The spoonerism task includes two subtests. In the first, listeners are required to drop the initial phoneme of a word and blend the resulting sequence with a phoneme or a cluster (“red with a [b] gives bed”). In the second, two words are presented and listeners are instructed to swap around the first sound of each word (“daisy log” gives “lazy dog”). Feedback is provided for the first three practice trials of each subtest. Each subtest contains ten test items, scored following the same procedure as the rhyme task.

Phonological Short Term Memory: Phonological short-term memory was assessed using the Nonword Repetition task (Gathercole and Baddeley, 1996). The test consists of 40 nonwords of 2 to 5 syllables in length (e.g , ‘rubid’, ‘sepretennial’) preceded by two practice items. The final score is the total number of nonwords that were repeated correctly.

Speech perception tests

Word perception in noise: To assess speech perception in noise, the participants completed two tasks in which they had to recognize naturally-produced words presented in background noise.

For the Words in Noise (WiN) test, 25 highly frequent monosyllabic words (e.g., “dog”, “cake”) were selected so as to correspond to an age of acquisition of no more than 4 years old (de Cara and Goswami, 2002). Items were presented in random order with multi-talker babble noise in the background fixed at 65 dB SPL (measured over a frequency range of 100-10Hz) and the Signal-to-Noise Ratio (SNR) varied by altering the level of the word. As some words are more robust than others in noise and thus able to tolerate lower SNRs, a preliminary calibration study was been performed in a previous study in order to determine a ‘correction factor’, uniquely specified for each word (Kunaratnam, 2003). Through this calibration, the SNR was adjusted to different values for different words to achieve a consistent baseline performance across words (Kunaratnam 2003). In the WiN test, the procedure started with an SNR of 12 dB and tracked 50% correct adaptively with a one-up one-down rule. The initial step-size was 6 dB, which decreased linearly over the first 4 reversals to 2 dB. The test ended after 10 reversals or 25 trials. Logistic regression was used to estimate the SRT (speech reception threshold – the SNR which leads to 50% words correct) from all trials run during the adaptive procedure.

The ‘Words in noise in connected speech’ (WiNiCS) task was modelled after the Coordinate Response Measure (Moore, 1981) as discussed in Brungart (2001). In this test, participants heard the following carrier phrase: “show the dog where the […] [...] is”, with the gaps filled by a colour and a number. In a trial, the six symbols on the screen were all the same number and differed only in colour (black, white, pink, blue, green, and red). Participants were instructed to click on the symbol that corresponded to the colour they heard. A three-up one-down adaptive procedure was used to vary SNR and so to track 79.4% correct trials. Unlike the ‘words in noise’ task described above, the total level of the output was fixed at 65 dB SPL. Therefore, as SNR decreased, the level of the speech decreased while the level of the babble increased. The first sentence was presented at an SNR of +20 dB, with an initial step-size of 10 dB which decreased linearly to 5 dB over the first 2 reversals. The test ended after a total of eight reversals or after 30 trials. The threshold for a 79.4% correct level was calculated from the mean of the reversals excluding the first two.

Categorical perception tasks: Phoneme categorization abilities were assessed by means of categorical perception tasks involving the identification and discrimination of a /pi/-/bi/ (‘pea’/’bee’) continuum in quiet and in noise.

Stimuli were generated by copy-synthesis using the cascade branch of a Klatt synthesizer (Klatt, 1980). The aim of copy-synthesis is to obtain a speech signal which is totally controllable but is also natural-sounding, as all parameters are copied from a specific utterance produced by a single speaker. Copy-synthesis was used as it has been suggested that the categorical perception deficits observed in children with SLI when tested with schematic synthetic speech do not generalize to tests using edited natural speech (Blomert and Mitterer, 2004). Initial values for fundamental and formant frequencies, vowel duration, and burst characteristics were measured from a natural [bi] token recorded by a female native British English speaker. The total syllable duration was 460 ms. For the first 4 ms, aspiration and friction amplitude were set at 74 and 70 dB respectively to produce a burst. Formant values (F1, F2, F3 and F4) were set at 365, 2000, 2600 and 4252 Hz respectively and reached 167, 2745, 3283 and 4119 Hz at the end of the syllable. The continuum was generated by delaying the onset of the voicing while concurrently increasing the aspiration duration, to obtain stimuli differing in Voice Onset Time (VOT) ranging from 0 ms for the [bi] endpoint to 60 ms at the [pi] endpoint of the continuum, in 1 ms steps (see Figure 1).

Pilot testing of the stimuli with 4 children and 4 adult monolingual English speakers indicated that the endpoint stimuli were convincing exemplars of the syllables /pi/ and /bi/. Responses to the labelling of a subset of 6 steps of the continuum differing in 10 ms VOT, exhibited the expected s-shaped categorisation function centred around 23 ms VOT, which is consistent with the location of the phoneme boundary in English (Abramson et al., 1967).

[Figure 1 about here]

Identification tasks: A two-alternative forced-choice task was used to assess category identification (AdaptID). Participants were instructed to identify the stimulus by clicking on a picture of a pea or a bee. Pictures were used rather than word labels in order to keep the test procedure consistent with that used in our study with children (Messaoud-Galusi, Hazan and Rosen, 2007). Stimuli were presented using an interleaved adaptive procedure as described in Ramus et al. (2003). The main advantage of an adaptive procedure is that trials are concentrated in the region most crucial for estimating the phoneme boundary and slope of the function, thus making an efficient use of a relatively small number of presentations. Another advantage is that the level of difficulty is consistent across participants as a particular level of performance (71% 'pea' or 'bee' responses) is tracked for each listener. Catch trials (continuum endpoints) were randomly interspersed 20% of the time so that participants would not hear an uninterrupted sequence of ambiguous stimuli. Two independent adaptive tracks were used. Each operated under identical rules except that they started at opposite ends of the continuum, and were designed to track 71% of 'bee' or 'pea' responses using a 2-down/1-up rule (Levitt, 1971). On any particular trial, the choice of track was made at random. The initial step-size was 10 ms, reducing linearly over the first 3 reversals to 3 ms. The initial track ascent/descent used a 1-down/1-up rule to move quickly into the region of interest, switching to the 2/1 rule after the first reversal. The interspersed endpoints also provided a measure of response consistency to ‘easy’ endpoint stimuli throughout the task. The task ended after 7 reversals on each track or a maximum of 50 trials.

For each listener and condition, responses to all test trials (i.e., excluding catch trials) were aggregated and logistic regression used to obtain a best-fit sigmoid function. Estimates of the slope and boundary were then obtained from the fitted coefficients. The boundary locates the point on the continuum at which ‘pea’ and ‘bee’ responses are equally probable, in other words the point at which the percept changes from one phonemic category to the other (the so-called phoneme boundary). The slope of the identification function is a measure that reflects the consistency with which the listener is categorizing the continuum. A shallower slope indicates a lower degree of consistency in the labelling of the continuum. The interspersed-endpoint trials were analysed separately and used as a measure of the level of attention maintained through the task.

The identification task was run in two conditions: in quiet (AdaptID-Q) and in noise (AdaptID-N). For the noise condition, multi-talker babble was played simultaneously with the word at an SNR of +6 dB. The total duration of the stimuli was 1000 ms with the noise starting about 315 ms before the beginning of the word. All other aspects of the stimuli were the same in quiet and in noise.

Discrimination tasks: Three different discrimination tasks were presented to each participant, using the same ‘pea’/’bee’ continuum: two adaptive discrimination tasks and a fixed discrimination task. A three-alternative forced-choice (3AFC) test procedure was used for all tasks. The task again was designed for use with children but could be run without problem with adults. Three frogs appeared on the screen with each ‘saying’ one of the stimuli from the continuum. Participants were told that two of the frogs would say something similar and one would say something different and were instructed to click on the frog that said something different. The ISI was set at 300 ms. A 3AFC procedure was preferred over a 2IAX procedure for the following reasons. First, chance level is lowered to 33%. Second, as discussed by Halliday and Bishop (2006), given that the odd stimulus can often be inferred by hearing the first two stimuli in the triplet, the third stimulus presented can provide further confirmation or refutation of the decision reached. Finally, previous studies with adult dyslexics have suggested that 2IAX procedures than lead to higher jnds, than 3AFC procedures at least for frequency discrimination tasks (France et al., 2002).

The ‘fixed reference discrimination task” (AdaptWC) was used to get a measure of just noticeable difference (jnd) within category. In this task, the standard stimulus for every test trial was the ‘pea’ endpoint of the continuum. The test started with the ‘bee’ endpoint as the comparison stimulus, which was presented for three trials, an easy discrimination for all. A 3-down/1-up adaptive procedure (Levitt, 1971) was used to choose the comparison stimulus so as to estimate the stimulus that could be discriminated from the standard 79.4% of the time. As for the identification task, a 1-down/1-up rule was used prior to the first reversal. Step-size varied throughout the test, from 12 ms VOT at the start, decreasing linearly over the first 3 reversals to 3 ms VOT. The task ended after 7 reversals on each track or a maximum of 50 trials. This test was done both in quiet (AdaptWC-Q) and in a background of babble noise of +6 dB SNR (AdaptWC-N). The jnd (just noticeable difference in VOT) was calculated by taking the mean of the final 4 reversals (i.e., when the minimum step-size had been reached). A jnd of less than 38 ms would typically indicate that the listener was able to discriminate differences within the [pi] category. This is because the jnd was with reference to the ‘pea’ endpoint (VOT=60 ms) and the mean phoneme boundary was at 22 ms VOT (60 ms – 22 ms= 38 ms VOT). As phoneme boundary points varied across listeners, an evaluation of whether each listener was discriminating within-category was made by comparing their discrimination threshold to their specific phoneme boundary point.

The ‘phoneme-boundary centred discrimination task’ (AdaptAC-Q) was used to get a measure of jnd across category in quiet. This task was essentially identical to the fixed reference discrimination task except that here, both the comparison and standard stimuli changed as the adaptive track proceeded, so as to remain centred on a phoneme boundary of 22.5 ms VOT (as determined in the pretesting of the continuum). Therefore, the standard ‘bee’ was initially set at 0 ms VOT and the comparison ‘pea’ at 45 ms VOT, resulting in jnds that were always across category and could lie between 1 and 45 ms. For example, the smallest jnd of 1 ms would be obtained in the final tokens centred on the phoneme boundary at 22.5 ms had VOTs of 23 ms and 22 ms. For both these tasks, larger jnds indicate poorer discrimination abilities.

In order to assess the consistency of performance in the phoneme discrimination task, and also to be able to compare our results more easily with previous studies of within- and across-category discrimination, a further discrimination test was presented in quiet using a fixed procedure. As for other discrimination tests, a three-alternative forced-choice (3AFC) test procedure was used, with the participants being asked to indicate which word was the ‘odd one out’ in the triplet presented. Six stimulus pairs were used: four within-category stimulus pairs (5-20 ms, 35-50 ms, 40-60 ms and 50-65 ms VOT) and two across-category pairs (20-35ms and 15-35 ms VOT). Each was presented 6 times in each of the following permutations of the stimuli A and B in a pair (ABA, ABB, AAB, BAB, BAA, BBA) giving 18 observations per stimulus pair (total:

108 observations). The proportion of correct responses was calculated for each pair, and mean scores were also calculated over the across-category pairs (FixedAC-Q) and within-category pairs (FixedWC-Q). Chance performance is 1/3, i.e. 33%.

Results

Phonological awareness and short-term memory

Mean scores obtained for the subtests of the PhAB task and Nonword Repetition tasks are presented in Table 2. The two participant groups did not differ significantly on the rhyme subtest of the PhAB but the DYS group performed significantly worse on the spoonerisms subtest and on the Nonword Repetition task, which assessed phonological short-term memory.

[Table 2 about here]

Perception of words in noise

The signal-to-noise threshold (dB SNR) for word intelligibility in babble noise was assessed in two tasks in which either high-frequency words were presented in isolation ('WiN' test) or a restricted set of colour categories had to be recognized within a sentence ('WiNiCS' test). Results are shown in Table 3. As expected, a higher level of noise could be tolerated in the WiNiCS given the highly-restricted vocabulary set, even in the face of the higher performance level demanded (79.4% correct tracked in WiNiCS versus 50% correct in WiN). The difference in thresholds between the DYS and AR groups did not reach significance for either of the two tests.

[Table 3 about here]

'Pea'/'bee' identification tasks

Figure 2 shows the summed data across participants in the AR and DYS groups for AdaptID-Q and AdaptID-N. These graphs show a high level of correct 'pea' and 'bee' identification for endpoint stimuli by both groups of listeners, despite claims in some previous studies of less consistent identification by dyslexic listeners in the endpoint regions of the continuum (e.g., Manis et al., 1997). To give a sense of individual performance on this task, estimated identification functions for individual participants (i.e., sigmoid curves from the fitting of individual data points) are presented in Figure 3. The slope measures were examined for AdaptID-Q and AdaptID-N, using data which excluded the interspersed endpoint presentations (see Table 3). As the distribution of slope measures was skewed, the log of the slope was used in order to obtain more symmetrical distributions. A repeated-measures ANOVA was carried out to evaluate the effect of participant group and test condition (quiet, in noise). Identification functions were sharper (reflecting better categorization) in quiet than in noise [F(1, 35)=63.38; p<0.001] but there was no significant group effect [F(1, 35)=1.54; p=.223] or group by condition interaction [F(1, 35)=0.076; p=.784], suggesting no evidence of poorer performance on this task by the DYS group either in quiet or in noise (see Table 3). The range of slope values was larger for the AR group in quiet (0.88 for AR group versus 0.77 for DYS group) but in noise, there was greater variance in the DYS group (range of 1.404 versus 1.295 for the AR group). A similar outcome was found for boundary measures: the mean phoneme boundary across all participants in the study (n=37) shifted from 22.0 ms VOT (s.d. 3.7) in quiet to 32.2 ms VOT (s.d. 13.4) in noise but there was no significant group effect [F(1, 35)=0.079; p=.781] or group by condition interaction [F(1, 35)=0.047; p=.829]. As the identification of catch trials (interspersed endpoints), which can be interpreted as an index of attention, was found to be significantly poorer for the DYS group than AR group in our study with children (Messaoud-Galusi, Hazan and Rosen, under review), this was also examined here. Both groups were at or near ceiling in quiet (100% for the DYS group and 98.6% for the AR group).

[Figures 2 and 3 about here]

‘Pea’/’bee’ discrimination tasks

First, consider performance on the adaptive tasks. The outcome measure of the AdaptWC task is the just-noticeable difference (jnd) in ms VOT from the endpoint ‘pea’ token (+60 ms VOT) which was the fixed reference. Within-category discrimination would be achieved if the jnd obtained in the AdaptWC task fell within the voiceless category for each listener. This was calculated in relation to the phoneme boundary measure obtained for that listener from the AdaptID-Q task. In quiet, 19/20 participants (95%) in the AR group, and 15/17 (88%) in the DYS group had thresholds that were within-category. In noise, only 8 participants in the AR group (40%) and 8 participants in the DYS group (47%) achieved within-category discrimination. The increase in standard deviation in the noisy condition indicates that there was a wider range in performance in both groups when noise was added (see Table 3). Because of the differences in variance across conditions, separate ANOVAs were carried out on the quiet and noisy conditions for AdaptWC to look at the effect of participant group: this was not significant in either condition (see Table 3) although there was a trend towards better performance by the AR group in clear. In the AdaptAC discrimination’ task, both adaptive tracks were varying so there was no fixed-reference acting as anchor. In this condition, the final threshold represents the across-category jnd. Again, the effect of participant group was not significant (see Table 3).

Second, performance on the fixed-procedure discrimination task was evaluated (see Figure 4). This test included four within-category pairs and two across-category pairs. First, one-way ANOVAs were carried out to see if discrimination varied across groups for any of the minimal pairs. Only discrimination for the 15-35 ms VOT ‘across-category’ pair was close to reaching significance even without correction for multiple comparisons [F(1,35)=4.005; p=.053], with better discrimination shown for the AR group. The scores for individual pairs were then aggregated to get mean across-category (FixedAC-Q) and within-category (FixedWC-Q) scores, and a repeated-measures ANOVA was used to evaluate the within-subject effect of type (within, across-category) and across-subject effect of group. The main effect of type was significant with a better discrimination of FixedAC-Q pairs [F(1, 35)=120.97; p<0.001]. The effect of group was also significant [F(1, 36)=4.544; p<0.05] but there was no significant group by type interaction [F(1, 35)=0.0739; p=.396]`. AR participants therefore showed better discrimination for both within-category and across-category pairs. Finally, the effect of step-size was examined for the two across-category pairs. A significantly higher score was obtained overall for the 20 ms-step pairs than for the 15 ms-step pairs [F(1, 35)=23.98; p<0.001] but there was no significant group by step-size interaction showing that the DYS group was not more affected by the step-size than the AR group.

This ’pea’-’bee’ discrimination test provides a rare opportunity to evaluate the effect of task procedure on speech perception tasks as within- and across-category discrimination were both evaluated using adaptive *and* fixed procedures with the same set of stimuli. For within-category discrimination, we compare performance for the 40 vs 60 ms pair in the FixedWC task with performance for the same interval in the AdaptWC task as estimated from the psychometric function. Recall that in the AdaptWC task, the fixed standard stimulus was always ‘pea’ at 60 ms, with the comparison stimulus changing as required by the adaptive procedure. For each set of data from a single adaptive test (representing performance by one listener), it is possible to plot performance as a function of the VOT of the comparison stimulus (the psychometric function). This will vary from chance (⅓) to perfect as the comparison stimulus varies from near 60 ms, to low VOT values at the ‘bee’ end of the continuum. Logistic regression (taking chance levels of performance into account) can then be used to obtain a best-fitting sigmoid curve to this psychometric function, and hence to estimate performance for the 40 vs 60 ms pair. Similarly, performance for the across-category pair in the FixedAC task (15 vs 35 ms VOT) can be compared to that estimated from the psychometric function in AdaptAC. These measures were calculated individually for each participant, and mean discrimination scores for each group are given in Table 4. A repeated-measures ANOVA was used to investigate the between-subject effect of group and within-subject effect of stimulus pair (40 vs 60 ms, 15 vs 35 ms) and test procedure (fixed, adaptive). The effect of listener group was not significant [F(1,35)= 3.195; p=0.08]. As expected, higher discrimination scores were obtained for the 15 vs 35 ms than for the 40 vs 60 ms pair [F(1,35)= 83.359; p<0.001]. There was a significant stimulus pair by test procedure interaction [F(1,35)= 9.603; p<0.005]: discrimination accuracy varied between the fixed and adaptive procedure for the within-category pair but not for the cross-category pair. The perception of within-category differences was therefore enhanced in both listener groups in the adaptive task in which there was a gradual reduction of the stimulus interval and a consistent reference stimulus (‘’pea’’ endpoint).

[Table 4 and Figure 4 about here]

Composite scores

As in Ramus et al. (2003), composite z-scores were then calculated to compare performance on reading tasks, phonological tasks, and speech perception in quiet and in noise. For each participant, a READING score was calculated by taking a mean of the z-scores for the TOWRE word and pseudoword subtasks, a PHONOLOGY score was calculated as the mean of the rhyme, spoonerism and nonword repetition z-scores, a QUIET score was calculated as the mean of the AdaptID-Q, FixedAC-Q, AdaptWC-Q and AdaptAC z-scores, and a NOISY score was calculated as a mean of the AdaptID-N, AdaptAC-N, WiN and WiNiCS z-scores (see Figure 5).The data was examined for outliers, defined as scores that were greater than two standard deviations below the mean for that group. Where outliers were found, statistical evaluations were carried out with and without outliers. It should be noted that no single individual was an outlier in more than one of these composite scores. The group effects are reported here with outliers included but any significant change in effect resulting from the removal of outliers is mentioned below. As expected, the AR and DYS groups differed in their READING score [F(1,35)=76.95); p<0.001]. They also differed in their PHONOLOGY scores [F(1,35)=9.027); p<0.01] and this group difference was even greater when one outlier per group was removed [F(1,33)=11.02; p<0.005]. The difference in the QUIET score just reached significance [F(1,35)=4.547; p<0.05], probably due to the poor performance on the fixed-step discrimination procedure by many individuals in the DYS group but the difference in the NOISY score did not [F(1,35)=0.859; p>0.05].

Correlations across the composite scores were then examined, for the data aggregated across the DYS and AR groups after outliers had been excluded. The READING score was significantly correlated with the PHONOLOGY score (r=.563; p=.001, N=34), QUIET score (r=.464; p=.006, N=34) and NOISY score (r=.363; p=.03, N=35),. There was a moderate correlation between the PHONOLOGY score and the QUIET (r=.433; p=.012, N=33) but not with the NOISY (r=.339, p=.05, N=34) scores. The QUIET and NOISY scores were correlated (r=.343; p=.04, N=34). When composite scores were examined separately for each group, none of the correlations reached significance.

[Figure 5 about here]

Individual differences

Several studies have suggested that only a subgroup of individuals with dyslexia may have speech perception difficulties (e.g., Lieberman et al., 1985, Adlard and Hazan, 1998; Ramus et al., 2003). It is therefore important to examine the performance of individual participants in both the DYS and AR groups to get a better sense of the proportion of individuals showing poor performance on speech perception tasks, even when group effects are not significant. A further rationale for this kind of analysis is to ascertain whether any participants are consistently poor at subtasks that are assessing a given processing ability, or whether poor performance appears to be more random, and therefore more likely to be due to reasons other than a perceptual deficit (Roach et al., 2004; Heath et al., 2006). Also, if poor categorization ability is likely to result in further perceptual difficulties when the speech signal is degraded, we expect to see that individuals showing poor performance on categorization tasks also show higher thresholds in the words in noise tasks.

Individual performance on the following eight tasks was examined. The WiN and WiNiCS tests both address the perception of words in noise, the AdaptID-Q and AdaptID-N both address phoneme categorization ability and the four scores from the ‘pea’/’bee’ discrimination tasks (AdaptWC-Q, AdaptWC-N, AdaptAC-Q, FixedAC-Q) all address the ability to discriminate subtle acoustic-phonetic changes using the same set of speech stimuli.

The method used to identify participants in each task who were performing below norm was as first described in Ramus et al (2003) and also used by Reid et al (2007). Average readers performing below 1.65 standard deviation of the mean for the AR group (i.e., 5th percentile) were removed, and the mean and standard deviation for the AR group was then recalculated. Any participant performing below 1.65 standard deviation of this ‘trimmed’ mean was considered to be performing ‘below norm’ for that task. This is a more stringent criterion than many studies, as, for example, Adlard and Hazan (1998) used a criterion of one standard deviation below the mean for average readers.

Overall, 2 adults from the DYS group (11.7%) and 9 from the AR group (45%) performed within norm on every one of the eight speech tasks and can be described as ‘good performers’ while 5 adults from the DYS group (29.4%) and 1 from the AR group (5%) performed ‘below norm’ on three or more of the eight speech tasks, and can be described as ‘poor performers’. The rest of the participants from the DYS group (58.8%) only fell below norm on one or two tasks. This analysis shows that despite the lack of a significant group effect for most of the speech tasks presented in the study, there is evidence of poorer performance by the DYS group, as only two participants in the DYS group are within norms on all speech tasks whereas half of the AR participants are.

Before assigning this poor performance to speech perception deficits, it is important to see whether those participants who showed poor performance did so consistently across tasks that were tapping the same level of processing (see Tables 5 and 6). As regards the perception of naturally-produced words in noise, none of the six participants in the DYS group performing below norm on one task (either WiN or WiNiCS) also showed below norm performance on the other. Only three participants in the AR group performed below norm for the WiN test and one for the WiNiCs test but there again none performed consistently badly for both tests. For the identification tasks, two participants within the DYS group were below norm for each of the two tests but none was below-norm for both AdaptID-Q and AdaptID-N. There were three or four poor performers for each of these tests within the AR group also, but only one participant was below-norm for both. The discrimination tasks are the most informative in terms of consistency in performance, as across and within-category discrimination in quiet for the 'pea'/'bee' continuum was tested using both fixed and adaptive tasks. 9/17 participants in the DYS group were below norm for the FixedAC-Q task and 6/17 for the AdaptAC-Q task, but only 3/17 were below norm for both. Within the AR group, 3/20 participants performed below norm for FixedAC-Q and 2 for AdaptAC-Q but only one of these participants was below norm for both. Finally, the individual data were examined to see whether any of the participants showed consistently poor performance for tasks presented in noise (fixed-reference discrimination in noise, identification in noise and two words in noise tasks). This was not the case for any of the AR or DYS participants.

Finally, the profile of the five poor-performers in the DYS group and one poor performer in the AR group were examined in more detail. The only test for which all six participants were below norm was the FixedAC-Q discrimination test. The other two tests on which they performed 'below norm' varied across the six individuals in this group. Their performance on nonverbal and verbal IQ, phonological short-term memory and the four composite scores was examined in more detail (See Table 7). Univariate ANOVAs were carried out to evaluate the effect of group ('poor performer', 'DYS good performer' or 'AR good performer') on these various scores. The three groups did not differ in terms of their non-verbal IQ but did in terms of their verbal IQ [F(2, 34)=4.329; p<0.05]. Post-hoc tests (Tukey's HSD) showed that the poor performer group had a significantly lower verbal IQ than the AR good performers but that DYS good performers did not differ significantly from either the poor performers or AR good performers. The same pattern of post-hoc analyses was obtained for the phonological STM task [F(2, 34)=6.600; p<0.005] and the PHONOLOGY composite score [F(2, 34)=7.189; p<0.005]. For the READING composite score, [F(2, 34)=31.750; p<0.0001], the DYS good performer and the poor performer group obtained lower scores than the AR good performer group. For the QUIET composite score, as expected, the effect of group was significant [F(2, 34)=8.876; p<0.001]; the poor performer group obtained lower scores than the AR and DYS good performer groups which did not differ from each other. This same pattern was also obtained for the NOISY score [F(2, 34)=5.734; p<0.01]. Overall, therefore, the DYS good performers, who were within norm on a majority of the speech tasks, achieved comparable scores to the AR group on non-verbal IQ, verbal IQ, phonological short-term memory and the PHONOLOGY composite score, and only differed in the READING composite score. The poor performer group, which included 5 DYS and 1 AR adult, however, showed poorer performance than the AR good performer group in terms of their verbal IQ, phonological STM and PHONOLOGY scores. The poor performers only differed significantly from the DYS good performers for the QUIET and NOISY speech scores.

[Tables 5, 6 and 7 here]

Discussion

This study tested adults with dyslexia and average readers on a range of speech perception tasks. Some of these tasks tapped the ability to identify speech sounds and discriminate subtle acoustic-phonetic differences within ’analytic’ tests in quiet and in noise (identification and discrimination skills). As the ability to discriminate a ‘pea’/’bee’ continuum was tested using two different methods (fixed or adaptive), it was possible to assess the consistency of any evidence of poor performance. Such consistency is key to attributing poor performance to a speech perception deficit rather than to other causes. Other tasks assessed the perception of naturally-produced words in noise. These more naturalistic tasks did not purely tap the use of acoustic-phonetic information, as listeners could also use lexical and phonotactic knowledge. It was still expected that any true deficit in phonemic categorization would lead to poor word perception in noise (Ramus, 2001).

The first aim of the study was to assess performance on categorical perception tasks, and to investigate whether the addition of noise in identification and discrimination tasks would lead to a greater decrease in performance for the DYS than for the AR group. A pattern of poor performance in quiet that worsens significantly in noise would suggest that phonemic categories in individuals with dyslexia may be underspecified and easily affected by further degradation of the signal. Overall, the group data revealed fewer across-group differences than many previous studies of speech perception abilities in adult dyslexics (e.g., Steffens et al., 1992; Schwippert and Koopmans-van Beinum, 1998; van Beinum et al, 2005). No significant differences between the AR and DYS groups were found in the steepness of the identification functions for a ‘pea’/’bee’ contrast both in quiet and in noise, nor for adaptive discrimination tasks for the same contrast. No group differences were found in the thresholds for the recognition of words in noise, whether the words were presented in isolation or in context. The only significant group difference was obtained for a fixed-step discrimination task for the same ‘pea’/’bee’ continuum, where significantly better discrimination of both within- and across-category pairs was shown for the AR group. There was therefore little evidence of consistently poorer categorization in the DYS group and it did not appear that their perception of speech was particularly affected by signal degradation.

Given that many studies have suggested that not all individuals with dyslexia may have speech perceptual processing difficulties, we need to consider whether a link between categorization and perception of speech in noise may be present at least for those few individuals who are performing poorly in the categorization and discrimination tasks, whether such individuals are dyslexic or average readers. However, as shown in the analysis of individual performance, there was no evidence that individuals who performed below norm on identification and discrimination tasks in quiet performed particularly poorly for the same tasks in noise or on the natural speech in noise tasks.

It is important to consider in what ways the speech perception tasks presented in this study differed from those in studies that did obtain group differences in the identification or discrimination of phonemic contrasts. This is not an easy comparison as studies differ in so many aspects of the stimuli and tasks used, and in the characteristics of the participant populations. In terms of stimuli, studies vary in the specific phonemic contrast used, whether the stimuli are synthesized or processed natural continua, and whether the target labels were lexical items or nonwords. Studies also differ in many aspects of task design, such as whether the task was fixed or adaptive, the step-size used and number of presentations.

One first source of variability is the phonetic contrast that was investigated. Many studies have tested contrasts in place of articulation (e.g. /ba/-da/), as these are cued by fast formant transitions, and thus were suspected to be particularly problematic for children with dyslexia or SLI (Tallal, 1980). The outcome of studies is inconsistent for these contrasts. For example, Steffens et al (1992) obtained significant group differences for a /ba/-/da/ contrast with adults but Ramus et al (2003) found no group differences in the identification of a ‘date’-‘gate’ contrast in their adult study. Results are equally inconsistent for studies investigating voicing contrasts. Ramus et al. (2003) obtained no significant group differences between dyslexic adults and average readers in the identification of a ‘coat’-’goat’ continuum, mirroring the result obtained here for a ‘pea’-‘bee’ contrast. However, Breier et al (2001) obtained a group difference in an identification task between dyslexic children and controls for a /ga/-/ka/ continuum with the greatest difference across groups being in the labeling of stimuli at the endpoints of the continuum. A similar group difference was found by Manis et al. (1997) for a ’path’-‘bath’ continuum, although they also point out that the majority of dyslexic children exhibited normal categorization, as only 7 out of 25 had abnormal identification functions. In French, poorer identification and discrimination was found by Bogliotti et al. (2008) for a /do/-/to/ contrast with children with dyslexia, and poorer discrimination was also obtained with a similar group for a /ga/-/ka/ continuum (Serniclaes et al., 2004).

Another source of variability is the method of stimulus construction used. Steffens et al. (1992) argued that marginally poorer perceptual performance in adults with dyslexia was likely only to be visible in situations in which linguistic context is absent, or which maximally stress phonetic perceptual abilities by removing cue redundancy, as occurs in rather schematic synthesized speech. This view that individuals with dyslexia may benefit from the redundancy of acoustic cues that is present in natural speech is supported by some studies showing better performance with natural than with synthetic speech tests (Lieberman et al., 1985; Masterson, Hazan and Wijayatilake, 1995). However, differences in categorization between DYS and AR groups have been obtained both for studies using fully-synthetic continua (e.g., Steffens et al., 1992; Breier et al., 2001) and those using computer-edited natural speech (e.g., Schwippert and Koopmans-van Beinum, 1998; van Beinum et al., 2005; Manis et al., 1997).

Finally, studies vary in the task procedures used in identification and discrimination tests. Fixed-step procedures present items that are fully-randomized and presented with equal frequency, while adaptive procedures track a specific level of performance for each individual, with the level of ambiguity of the stimuli increasing as the task progresses, at least in the initial stages of the test (apart from the catch trials). Given the suggestion that individuals with dyslexia have poor attention and short-term memory, it is conceivable that certain aspects of these procedures may affect performance. The comparison of within-category discrimination across the fixed and adaptive tests showed the degree to which performance could be affected by specific aspects of the test procedure. In this case, although the synthetic stimulus continuum and the 3IFC procedure used were the same across the two tests, within-category discrimination was better in both groups of participants for the adaptive procedure (AdaptWC), which used a fixed reference, tracked a specific level of accuracy and where the physical difference between the stimuli reduced during the test. When fixed-step discrimination procedures are used, a combination of task difficulty, longer test duration, and perceptual abilities within the lower range of a normal distribution could conceivably lead to poorer performance. For identification tasks, it could be argued that fixed procedures, which present 'easy' trials (e.g. tokens from the endpoint regions of the continuum) distributed throughout the test rather than at the beginning of the task, could be less difficult than adaptive tasks which focus presentations in the more ambiguous region of the continuum. However, adaptive procedures counter this by typically interspersing endpoint stimuli 20% of the time, and tend to achieve good estimates of slope and phoneme boundary measures with a smaller number of presentations so make a more efficient use of limited attention spans. A comparison of procedures in past studies is not very informative as studies vary in many aspects other than the task procedure. Most studies with dyslexic adults or children have used fixed-step procedures for their identification tasks. To our knowledge, the exceptions are studies by Adlard and Hazan (1998) with children and Ramus (2003) with adults which included adaptive consonant place and voicing identification tasks in their test battery. The Adlard and Hazan (1998) study was not fully adaptive as the stimulus continuum only contained six stimuli, but presentations were focused in the phoneme boundary region. As in our study, neither of these two studies obtained significant group differences in the slopes of the identification functions for any of the contrasts. However, all three studies also differ from other studies in using copy-synthesized stimuli, in which the syntheses are carefully matched to a natural utterance, rather than either stylised syntheses or natural edited speech. Both factors could therefore have led to improved performance in the dyslexic group.

Task-related issues may also partly explain the discrepancy between our results for tests presented in noise, which failed to show any group differences for phoneme identification, phoneme discrimination or word identification tasks, and previous studies which suggested that children or adults with dyslexia perform particularly poorly in noisy conditions. Ziegler et al. (in press) found no differences between DYS and AR children in the identification of naturally-produced VCV tokens in quiet, with both groups showing ceiling effects, but they obtained significant group differences at all noise levels. The seven separate identification tests, presented using a 16 alternative forced-choice fixed-procedure, were counterbalanced across listeners but the ‘silence’ condition was always presented first. Fatigue and lapses of concentration could therefore conceivably have affected the scores for noisy conditions more than the scores in the ‘silent’ condition. No information was provided on the performance of individual participants. In their study with dyslexic adults, Ramirez and Mann (2005) also showed evidence of a greater decrease in consonant accuracy in the DYS relative to AR group for nonsense CV syllables. Here, task difficulty was potentially increased by the use of a fixed procedure and of a full randomization within a single test of audio-only, visual-only and audiovisual stimuli presented both in silence and in different noise conditions. Therefore, although both studies may genuinely reflect difficulties with speech in noise, alternative explanations based on task-related factors are also plausible.

Our results therefore suggest the following picture. First, any claim of a causal link between dyslexia and speech perception difficulties seems questionable in the light of so many studies that show a majority of individuals with dyslexia to be within norms for speech perception tasks despite poor phonological processing (see also Ramus, White and Frith, 2006). A weaker proposal is that only a subset of children or adults with dyslexia with poor phonological processing may have speech perceptual deficits (e.g. Adlard and Hazan, 1998; Manis and Keating, 2005). Under this view, whether group differences are significant or not would depend on the proportion of individuals in the dyslexic cohort that happens to have a speech perception deficit, thus explaining the inconsistency found across studies. A final view may be that the poor performance shown by some individuals with dyslexia on speech perception tasks may not be due to a significant speech perceptual deficit. In our study, a comparison of performance on the fixed-step and adaptive ‘pea’/’bee’ discrimination tasks, and on the identification of the same contrast in quiet and in noise, can inform about whether poor performance in individual participants is consistent across related perceptual tasks. The lack of consistent poor performance across tasks for ‘poor performers’ provides little support for a specific speech perceptual deficit in dyslexic adults. On the basis of our data, it is not possible to totally discount the possibility that certain individuals with dyslexia have poor speech perceptual abilities but the alternative explanation of poor performance being due to non-sensory factors also seems plausible. Sutcliffe, Bishop, Houghton and Taylor (2006), for example, found that significant relationships between frequency discrimination performance and measures of language and reading ‘were abolished when comorbid attentional difficulties were taken into account’. On the other hand, a study that compared the performance of dyslexic children with and without ADHD on auditory perception tasks suggests that attention alone (at least as expressed in ADHD) cannot fully account for poor performance on these tasks (Breier et al., 2001). It may also be that in some individuals, speech perceptual acuity in the lower end of the normal range combined with task-related and other non-sensory factors such as lapses in concentration (e.g., Davis et al, 2001) may be sufficient to lead to ‘deviant’ performance on some tests in the battery. Further studies could attempt to elucidate this question by including other measures of attention and short-term memory, and tasks that tap phoneme categorization indirectly. These could include tests that evaluate the impact of varying degrees of within- and across-speaker variability on consonant discrimination and identification in naturally-produced words.

Acknowledgments

This study was funded by the Wellcome Trust (076499/Z/05/Z). The authors thank Mike Coleman who designed the testing software used in this study, Steve Nevard for technical support and Sam Eaton-Rosen for producing Figure 1.

References

Adlard, A. and Hazan, V. (1998).Speech perception in children with specific reading difficulties (dyslexia).*Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology*,*51*,153-177.

Blomert, L. and Mitterer, H. (2004). The fragile nature of the speech-perception deficit in dyslexia: Natural vs. synthetic speech. *Brain & Language*, 89, 21-26.

Blomert, L., Mitterer, H., Paffen, C. (2004). In search of the Auditory, Phonetic, and/or Phonological Problems in Dyslexia: Context effects in Speech Perception. *Journal of Speech, Language, & Hearing Research,* 47, 1030-1047

Boada, R. and Pennington, B. F. (2006).Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology*, *95*, 153-193.

Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., Sprenger-Charolles, L. (2008). Discrimination of speech sounds by dyslexic children: Comparisons with chronological age and reading level controls. *Journal of Experimental Child Psychology*, 101, 137-155.

Breier, J. I., Gray, L., Fletcher, J. M., Diehl, R. L., Klaas, P., Foorman, B. R., Molis, M. R.(2001).Perception of voice and tone onset time continua in children with dyslexia with and without attention deficit/hyperactivity disorder. *Journal of Experimental Child Psychology*, *80*, 245-270.

Brungart, D.S. (2001) Evaluation of speech intelligibility with the coordinate response measure. *Journal of the Acoustical Society of America*, *109*, 2276-2279.

Cornelissen, P. L., Hansen, P. C., Bradley, L., Stein, J. F. (1996). Analysis of perceptual confusions between nine sets of consonant-vowel sounds in normal and dyslexic adults. *Cognition*, *59*, 275-306.

Davis, C., Castles, A., McAnally, K.,Gray, J. (2001). Lapses of concentration and dyslexic performance on the Ternus Task. *Cognition*, 81, B21-31.

De Cara, B., and Goswami, U. (2002). Statistical Analysis of Similarity Relations among Spoken Words: Evidence for the Special Status of Rimes in English. *Behavioural Research Methods and Instrumentation,* 34*,* 416-423.

Dunn, L.l.M., Dunn, L.M., Whetton, C., Burley, J. (1997). British Picture Vocabulary Scale 2nd edition (BPVS-II). Windsor, Berks: NFER-Nelson.

Elbro, C, Nielsen, I., Petersen, D.K. (1994). Dyslexia in adults – evidence for deficits in non-word reading and in the phonological representation of lexical items. Annals of Dyslexia, 44, 205-226.

France, S., Rosner, B., Hansen, P., Calvin, C., Talcott, J. B., Richardson, A., & Stein, J. (2002). Auditory frequency discrimination in adult developmental dyslexics. Perception and Psychophysics, 64, 169–179.

Gathercole, S. E., and Baddeley, A. D. (1996). The Children's Test of Nonword Repetition. London: Psychological Corporation.

Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K. K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology*, 32, 401–424.

Halliday, L.F. and Bishop, D.V.M. (2006) Auditory frequency discrimination in children with dyslexia. Journal of Research in Reading, 29, 213-228.

Heath, S. M., Bishop, D. V. M., Hogben, J. H., Roach, N. W. (2006).Psychophysical indices of perceptual functioning in dyslexia: A psychometric analysis. *Cognitive Neuropsychology*, *23*, 905-929.

Heath, S. M. and Hogben, J. H.(2004). The reliability and validity of tasks measuring perception of rapid sequences in children with dyslexia. *Journal of Child Psychology and Psychiatry*, *45*, 1275-1287.

Joanisse, M., Manis, F., Keating, P., Seidenberg, M. (2000). Language Deficits in Dyslexic Children: Speech Perception, Phonology and Morphology. Journal of Experimental Child Psychology, 77, 30-60.

Klatt D. (1980). Software for a Cascade/Parallel Formant Synthesizer. *Journal of the Acoustical Society of America*, 67, 971-995.

Kunaratnam, N. (2003) Speech in Babble verses Speech in Noise in Normal Subjects and in Patients with Neurological Lesions. MSc(Audiological Medicine) Thesis, UCL, London.

Laasonen, M., Service, E., Virsu, V. (2002). Crossmodal temporal order and processing acuity in developmentally dyslexic young adults. *Brain and Language*, *80*, 340-354.

Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *Journal of the Acoustical Society of America*, *49*, 467-477.

Lieberman, P., Meskill, R.H., Chatillon, M., Schupack, H. (1985). Phonetic speech perception deficits in dyslexia. *Journal of Speech and Hearing Research, 28*, 480-486.

Maassen, B., Groenen, P., Crul, T., Assman-Hulsmans, C., Gabreels, F. (2001).Identification and discrimination of voicing and place-of-articulation in developmental dyslexia. *Clinical Linguistics & Phonetics*, *15*, 319-339.

McArthur, G. M. and Hogben, J. H. (2001). Auditory backward recognition masking in children with a specific language impairment and children with a specific reading disability. *Journal of the Acoustical Society of America*, 109, 1092–1100.

McArthur, G.M., Ellis, D., Atkinson, C., Coltheart, M. (2008). Auditory processing deficits in children with reading and language impairments: Can they (and should they) be treated? *Cognition*, 107, 946–977.

Manis, F. R., & Keating, P. (2005) Speech perception in dyslexic children with and without language impairments. In H. W. Catts & A. G. Kamhi (Eds.). *The connections between language and reading disabilities.* (pp. 77-99). Mahwah, NJ: Erlbaum.

Manis, F. R., McBride Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., Petersen, A.(1997).Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology*, *66*, 211-235.

Masterson, J., Hazan, V., Wijayatilake, L. (1995).Phonemic Processing Problems in Developmental Phonological Dyslexia. *Cognitive Neuropsychology*, *12*, 233-259.

Messaoud-Galusi, S., Hazan, V., Rosen, S. (2007). Clarifying the speech perception deficit in dyslexic children. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrueken, Germany, 6-10 August 2007, Saarbrueken, Germany, 697-700

Mody, M., Studdert-Kennedy, M., Brady, S. (1997).Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology*, *64*, 199-231.

Moore, D.R., Ferguson, M.A., Halliday, L.F., Riley, A. (2008). Frequency discrimination in children: Perception, learning and attention**.** *Hearing Research*, 238, 147-154.

Moore, T. (1981). Voice communication jamming research. *AGARD Conference proceedings* 331: Aural Communication in aviation, 2:1-2:6.

Ramirez, J. and Mann V. (2005). Using auditory-visual speech to probe the basis of noise-impaired consonant–vowel perception in dyslexia and auditory neuropathy. *Journal of the Acoustical Society of America*, 118, 1122-1133.

Ramus, F. (2001).Outstanding questions about phonological processing in dyslexia. *Dyslexia*, *7*, 197–216.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., Frith, U. (2003).Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, *126*, 841-865.

Ramus, F., White, S., Frith, U. (2006). Weighing the evidence between competing theories of dyslexia. *Developmental Science, 9,* 265-269.

Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of Experimental Child Psychology*, 48, 270–292.

Reid, A.A., Szczerbinski, M., Iskierka-Kasperek, E., Hansen, P. (2007). Cognitive profiles of adult developmental dyslexics: Theoretical implications. *Dyslexia, 13,* 1-24.

Roach, N. W., Edwards, V. T., Hogben, J. H. (2004). The tale is in the tail: An alternative hypothesis for psychophysical performance variability in dyslexia. *Perception*, *33*, 817-830.

Robertson, E.K., Joanisse, M.F., Desroches, A.S., Ng, S. (in press). Categorical speech perception deficits distinguish language and reading impairments in children. *Developmental Science.*

Schwippert, C.E. & Koopmans-van Beinum, F.J. (1998). Phoneme boundary perception in dyslexic and normal-reading adults. *Proceedings of the Institute of Phonetic Sciences Amsterdam*, 22, 125-134.

Serniclaes, W. I., Van Heghe, S., Mousty, P., Carre, R., Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology*, *87*, 336-361.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., Shankweiler, D. P., Liberman, A. M., Skudlarski, P., Fletcher, J. M., Katz, L.,

Marchione, K. E., Lacadie, C., Gatenby, C., Gore, J. C. (1998).Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America*,*95*,2636-2641.

Skottun, B. C. and Skoyles, J. (2007). Dyslexia: Sensory deficits or inattention? *Perception*, *36*, 1084-1088.

Snowling, M.J. (2000) *Dyslexia (2nd Edition)*, Blackwell Publishers: Oxford.

Snowling, M. J., Nation, K., Moxham, P., Gallagher, A., Frith, U. (1997). Phonological processing skills of dyslexic students in higher education: a preliminary report. *Journal of Research in Reading*, 20, 31-41.

Steffens, M. L., Eilers, R. E., Grossglenn, K., Jallad, B. (1992). Speech-Perception in Adult Subjects with Familial Dyslexia. *Journal of Speech and Hearing Research*, *35*, 192-200.

Sutcliffe, P., Bishop, D., Houghton, S., & Taylor, M. (2006). Effect of attentional state on frequency discrimination: A comparison of children with ADHD on and off medication. *Journal of Speech, Language and Hearing Research*, *49*, 1072-1084

Szenkovits, G. and Ramus, F. (2005).Exploring dyslexics' phonological deficit I: Lexical vs sub-lexical and input vs output processes. *Dyslexia*, *11*, 253-268.

Talcott, J. B., Hansen, P. C., Willis-Owen, C., McKinnell, I. W., Richardson, A. J., Stein, J. F. (1998). Visual magnocellular impairment in adult developmental dyslexics. *Neuro-ophthalmology*, 20, 187–201.

Tallal, P. (1980). Auditory Temporal Perception, Phonics, and Reading Disabilities in Children. *Brain and Language*, *9*, 182-198.

Torgesen, J.K., Wagner, R., Rashotte, C. (1997). Test of Word Reading Efficiency. Austin, TX: Pro-Ed.

van Beinum, F.J., Schwippert, C.E., Been, P.H., van Leeuwen, T.H., Kuijpers, C.L.T. (2005). Development and application of a /bAk/-/dAk/ continuum for testing auditory perception within the Dutch longitudinal dyslexia study. *Speech Communication*, 47, 124-142.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Revised*. San Antonio, TX: The Psychological Corporation.

Werker, J., & Tees, R. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology*, 41, 48-61.

Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., Stein, J. F., Green, G. G. R. (1998). Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Current Biology*, *8*, 791-797.

Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F. X., Perry, C. (2008). Developmental dyslexia and the dual route model of reading: simulating individual differences and subtypes. *Cognition, 107,* 151–178.

Ziegler, J. C., Pech-Georgel, C., George, F., Lorenzi, C. (in press). Speech-perception-in-noise deficits in dyslexia. *Developmental Science.*

Figure captions

Figure 1: Waveforms and spectrograms of the ‘pea’ (VOT: 60 ms) and ‘bee’ (VOT: 0 ms) endpoints of the speech continuum used to measure categorical perception in quiet and in noise

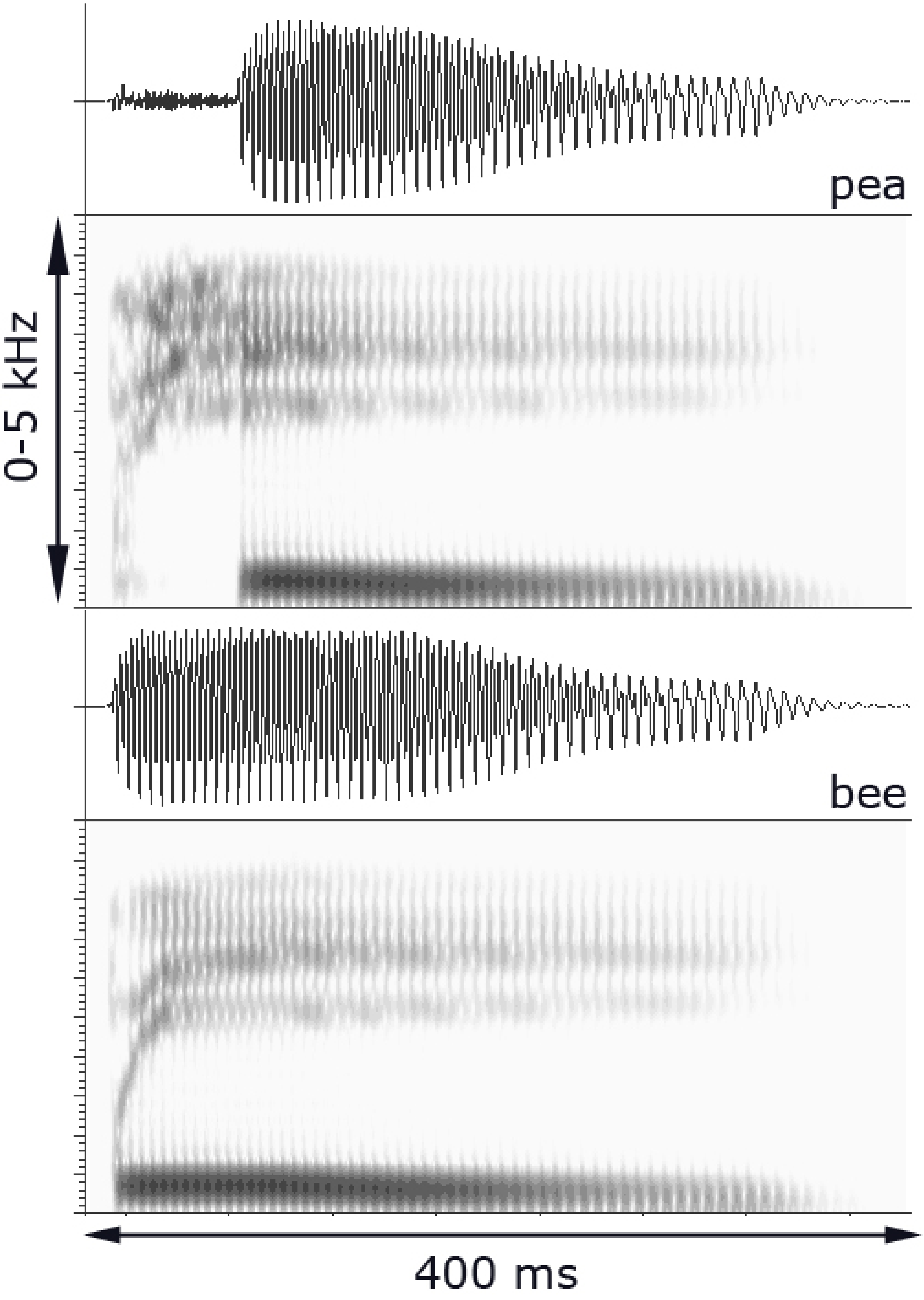
Figure 2: Summed data across participants in the AR and DYS groups for the AdaptID-Q and AdaptID-N tests. The size of the circle at a particular step is related to the total number of presentations at that step. Logistic regression was then used to obtain a best-fit sigmoid function for each set of data.

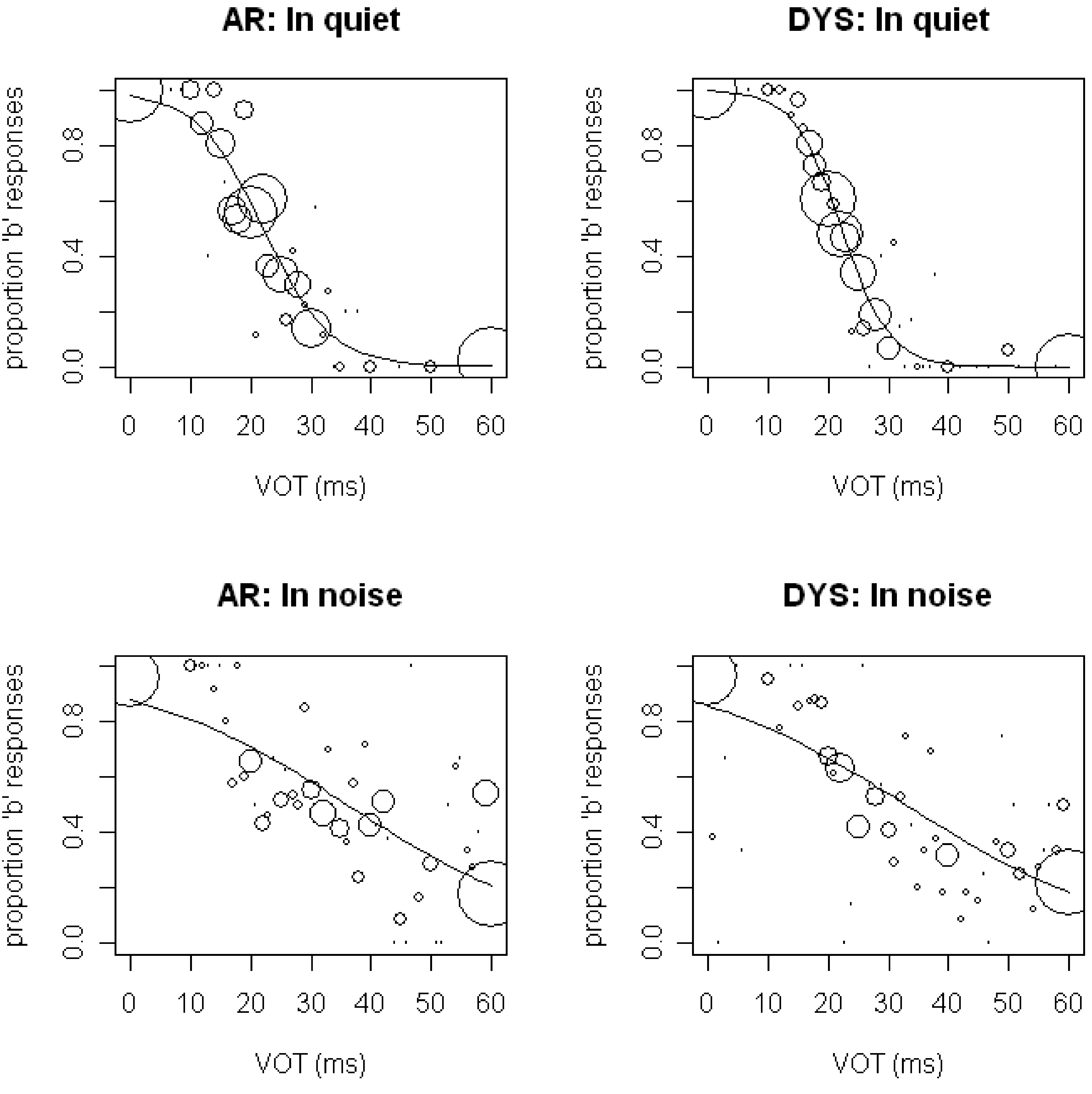
Figure 3: Individual identification functions for ‘pea’/’bee’ identification in quiet (AdaptID-Q) and in noise (AdaptID-N) for participants in the DYS and AR groups. The curves were extrapolated from the individual data points using a best-fit sigmoid function.

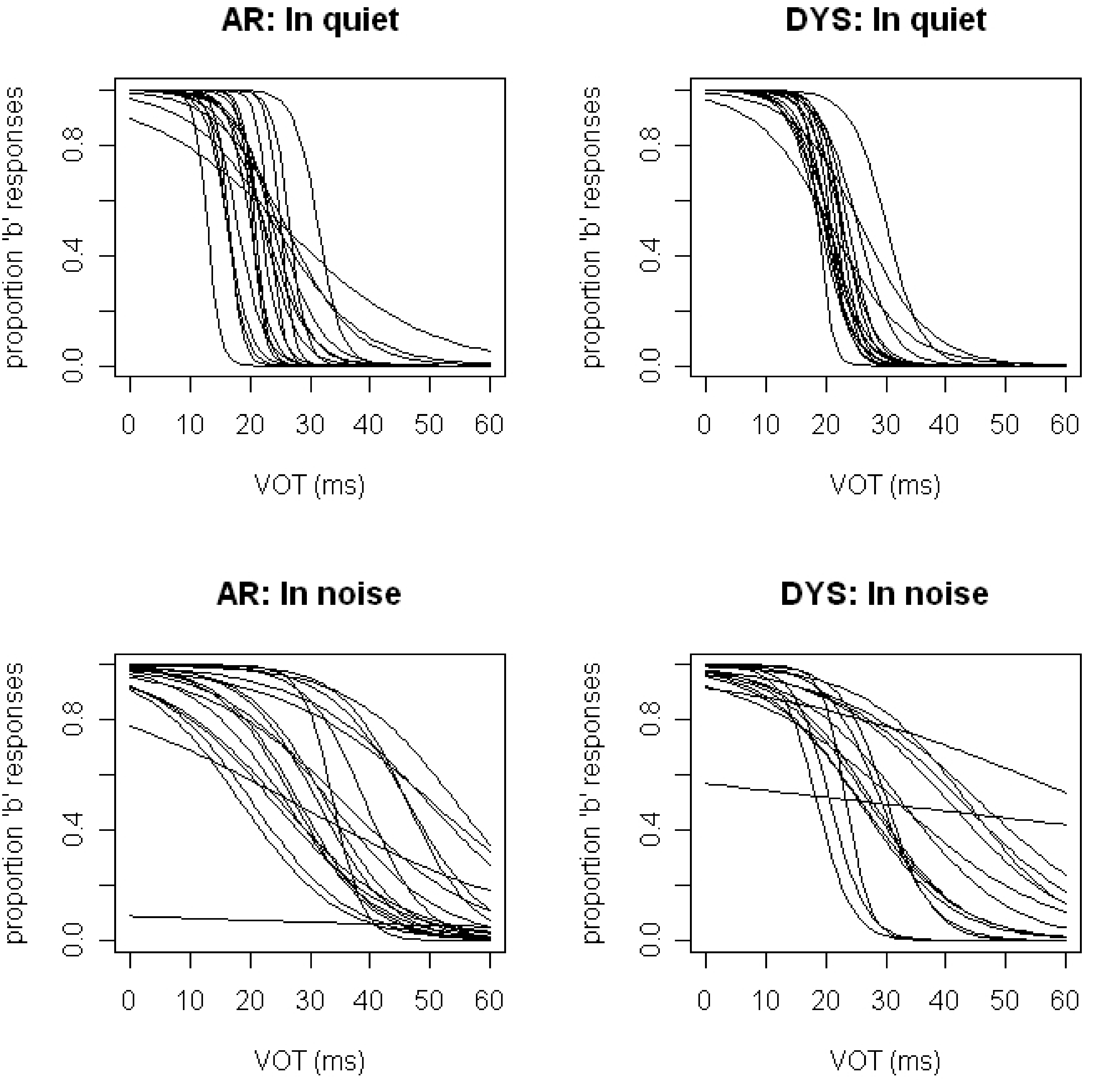
Figure 4: Box-plots showing correct discrimination scores for the AR (white boxes) and DYS (grey boxes) groups on the six stimulus-pairs, which are labeled as either within-category (WC) or across-category (AC) pairs. Boxplots display the first and third quartiles (edges of the box), median (horizontal line) and minimum and maximum values that are not outliers (whiskers). Outliers, displayed as circles, are cases with values between 1.5 and 3 box-lengths from the quartiles. Extremes, displayed as asterisks, are cases with values greater than 3 box-lengths from the quartiles.

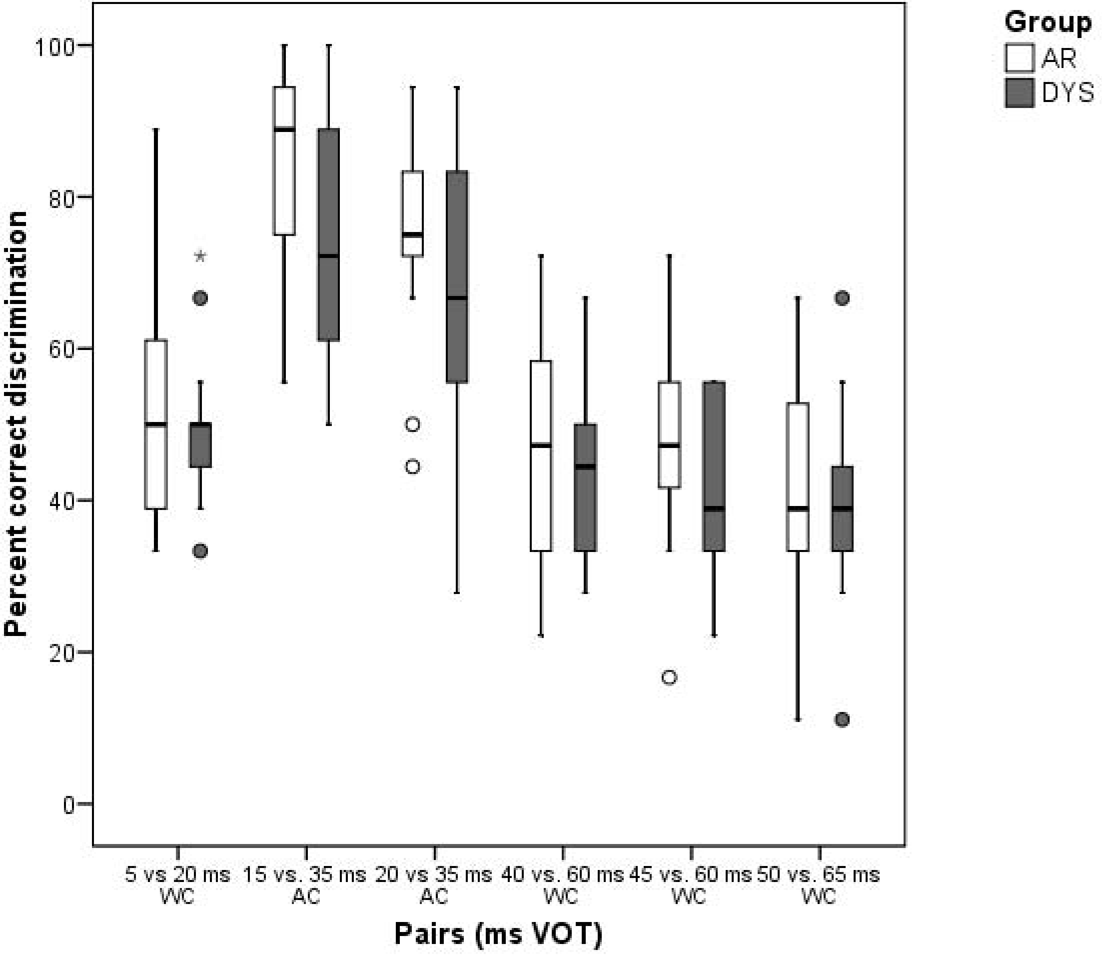
Figure 5: Box-plots showing the composite z-scores for READING, PHONOLOGY, QUIET (speech perception tests in quiet) and NOISY (speech perception in noise) skills for the AR and DYS groups.

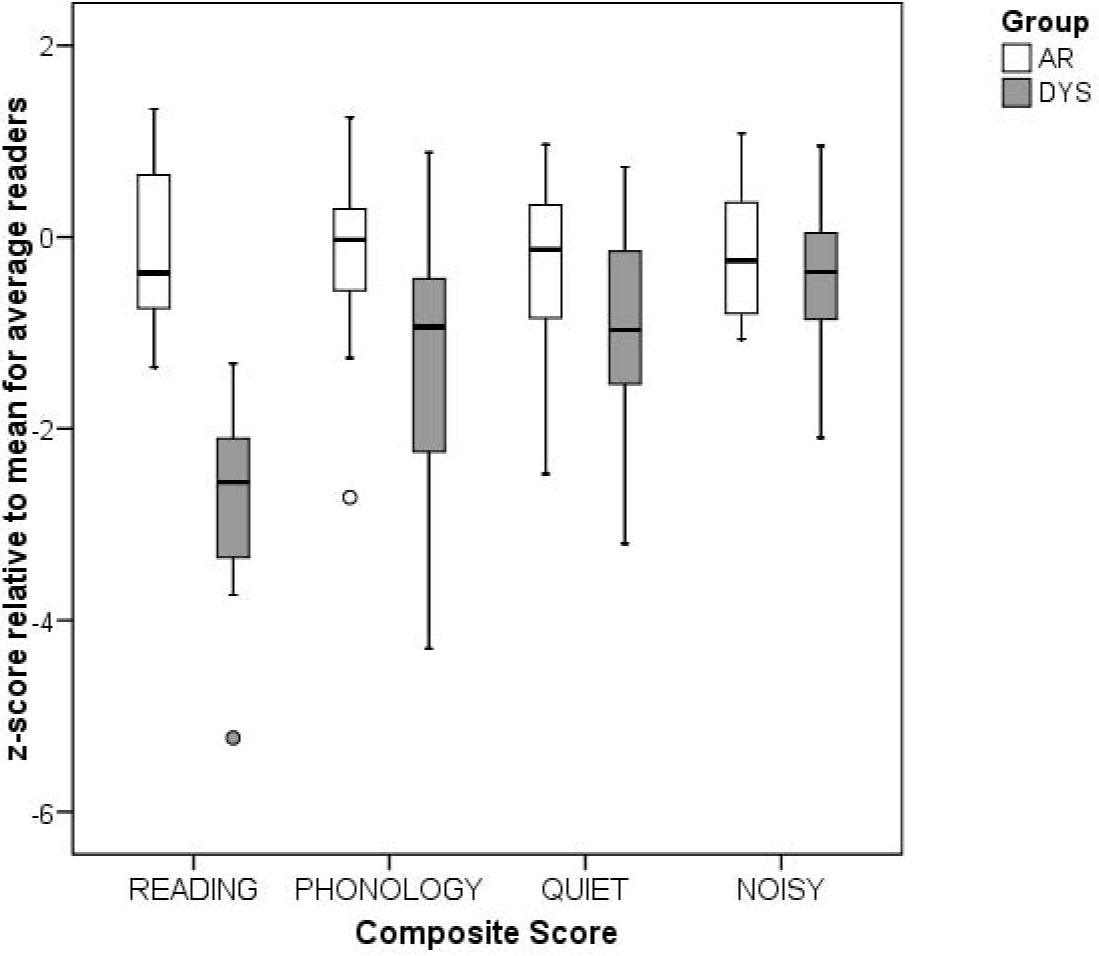
Figure 6: Individual data points for participants in the AR and DYS groups for the composite z-scores for READING, PHONOLOGY, QUIET (speech perception tests in quiet) and NOISY (speech perception in noise) skills. These z-scores are calculated relative to the means obtained for the AR group. The horizontal line represents the point at which scores were 1.65 standard deviations below the mean for the AR group. Individual performance below that point is considered ‘below norm’.

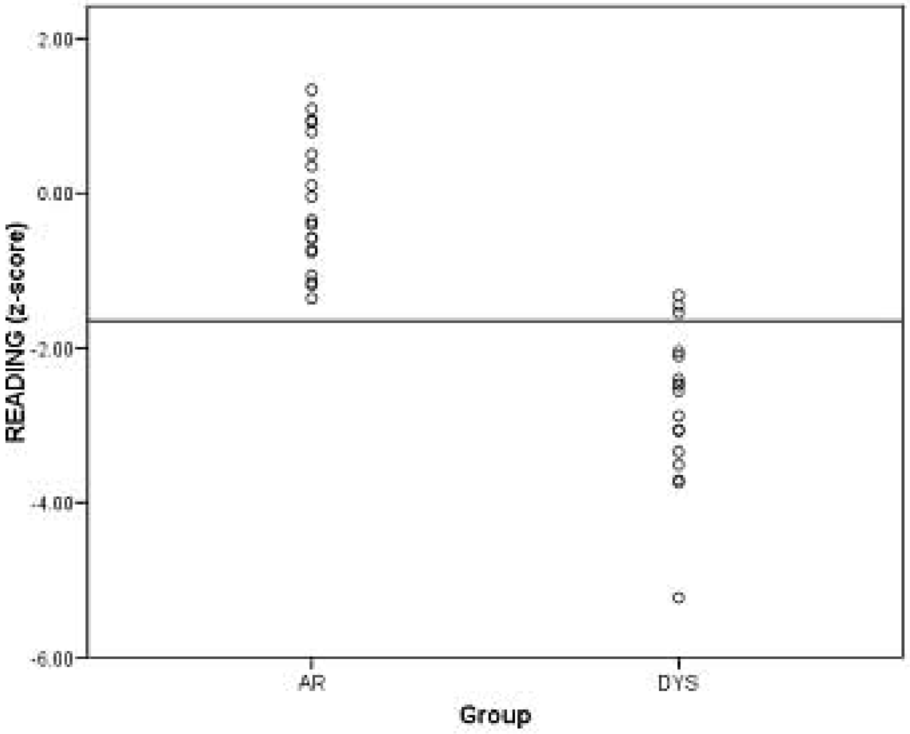

pea
0-5 kHz
bee
400 ms

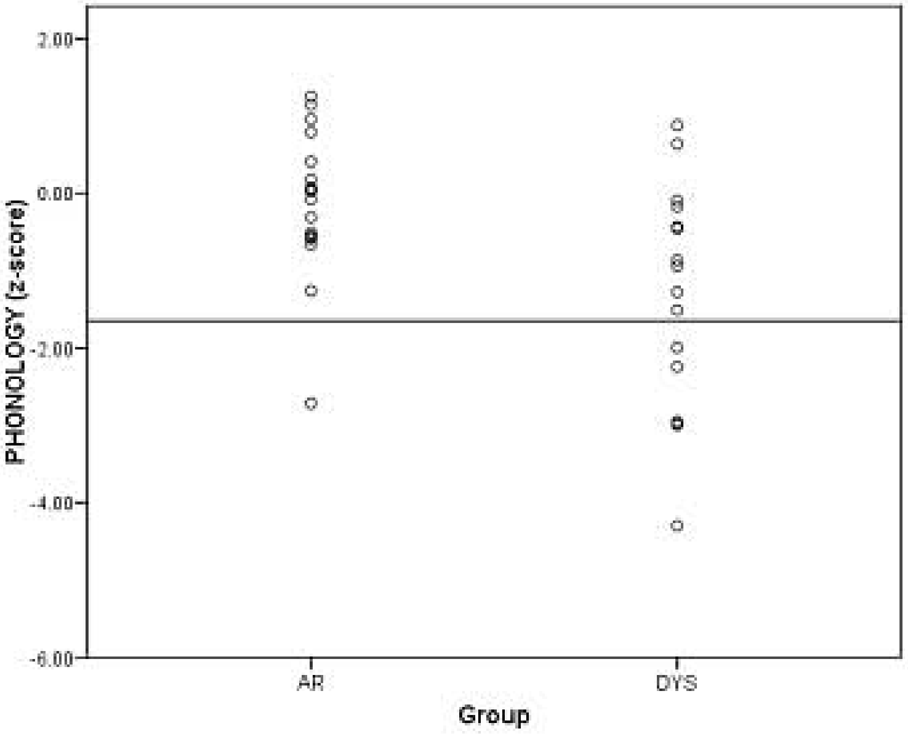

AR: In quiet
DYS: In quiet
AR: In noise
DYS: In noise
proportion 'b' responses
0.0
0.4
0.8
VOT (ms)
0
10
20
30
40
50
60

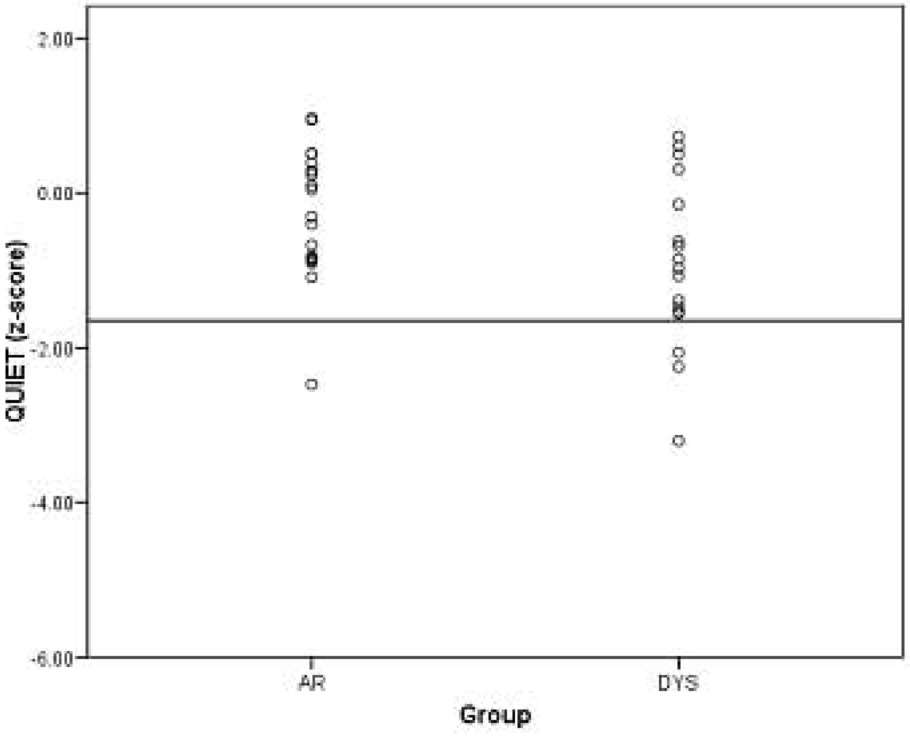

AR: In quiet
DYS: In quiet
AR: In noise
DYS: In noise
proportion 'b' responses
0.0
0.4
0.8
0
10
20
30
40
50
60
VOT (ms)


Group
AR
DYS
Percent correct discrimination
100
80
60
40
20
0
5 vs 20 ms
WC
15 vs. 35 ms
AC
20 vs 35 ms
AC
40 vs. 60 ms
WC
45 vs. 60 ms
WC
50 vs. 65 ms
WC
Pairs (ms VOT)


Group
AR
DYS
z-score relative to mean for average readers
2
0
-2
-4
-6
READING
PHONOLOGY
QUIET
NOISY
Composite Score


READING (z-score)
2.00
0.00
-2.00
-4.00
-6.00
AR
DYS
Group


PHONOLOGY (z-score)
2.00
0.00
-2.00
-4.00
-6.00
AR
DYS
Group


QUIET (z-score)
2.00
0.00
-2.00
-4.00
-6.00
AR
DYS
Group

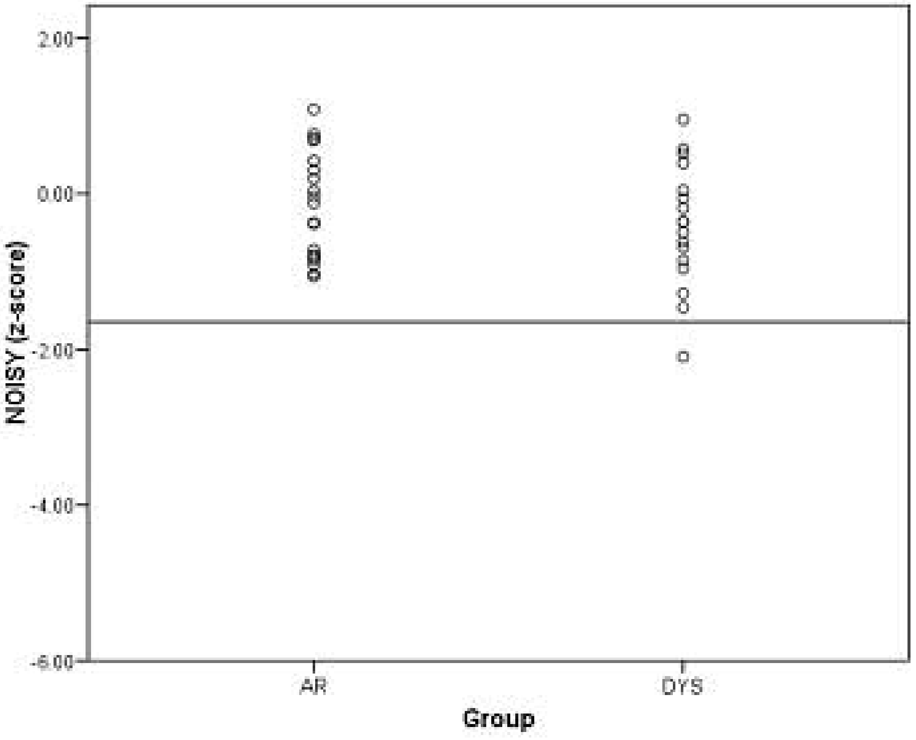

NOISY (z-score)
2.00
0.00
-2.00
-4.00
-6.00
AR
DYS
Group

Table 1: *Mean scores (and standard deviations) for the DYS and AR groups for: age, grammar, non-verbal IQ, verbal IQ and reading assessments. The last two columns present the results of independent-samples t-tests (with 'group' as a between-subject factor).*

Group	**DYS (N=17)**		**AR (N=20)**			
	Mean	**s.d.**	**Mean**	**s.d.**	**t**	**sig.**
Age (months)	273.9	33.7	281.0	42.0	.56	.58
Grammar (TROG)	98.7	7.5	100.3	5.8	.71	.49
Non-verbal IQ (WAIS-III)	116.2	16.0	112.8	12.5	-.73	.47
Verbal IQ (BPVS)	128.7	20.0	145.5	16.8	2.78	.01
Word Reading (TOWRE)	77.9	10.7	100.7	10.3	6.57	.00
Nonword Reading (TOWRE)	79.6	10.2	105.3	10.1	7.70	.00
Reading (aggregate score)	74.5	10.9	103.7	9.2	8.85	.00

Table 2: *Mean standardized scores and standard deviation measures for the average reader (AR) and dyslexic (DYS) groups on the Rhyme and Spoonerisms subtests of the PhAB, which assess phonological awareness, and on the phonological short-term memory (nonword repetition) test. The last two columns present the results of independent-samples t-tests (with 'group' as a between-subject factor).*

Group	**AR (N=20)**		**DYS (N=17)**			
	Mean	**s.d.**	**Mean**	**s.d.**	**t**	**sig.**
Rhyme (PhAB)	88.8	*7.0*	87.7	*7.7*	0.70	.49
Spoonerisms (PhAB)	86.5	*10.9*	77.4	*13.4*	2.30	.03
Phonological STM (nonword repetition)	93.8	*5.5*	83.5	*13.6*	2.90	.01

Table 3: *Mean scores and standard deviation measures for the AR and DYS groups for all the speech tests presented using an adaptive procedure. Scores for WiN represent dB SNR values at the speech reception threshold (SRT), AdaptID measures are the slope value for test items only (catch trials excluded) and measures for the PEA-BEE discrimination scores (AdaptAC-Q, AdaptWC-Q, AdaptWC-N) are the jnd in ms VOT. The last two columns present the results of independent-samples t-tests (with 'group' as a between-subject factor).T-tests for the AdaptID tests were carried out on the logs of the slope values.*

Group	**AR (N=20)**		**DYS (N=17)**			
	Mean	**s.d.**	**Mean**	**s.d.**	**t**	**sig**
WiN	-5.99	*1.50*	-5.21	*1.59*	-1.53	.14
WiNiCS	-8.27	*1.56*	-8.49	*2.05*	1.45	.16
AdaptID-Q	-0.59	*0.28*	-0.44	*0.18*	1.39	.17
AdaptID-N	-0.18	*0.11*	-0.17	*0.13*	0.64	.53
AdaptAC-Q	16.25	*9.68*	19.82	*11.70*	-1.02	.32
AdaptWC-Q	22.85	7.75	27.12	*9.17*	-1.53	.13
AdaptWC-N	32.30	*14.40*	32.70	*12.10*	-0.08	.93

Table 4: *Mean discrimination scores (% correct) for the across-category 15-35 ms VOT pair and within-category 40-60 ms VOT pair in the fixed and adaptive tests. For the adaptive tests, the discrimination score for the 15-35 ms pair was estimated from the psychometric function for AdaptAC and the score for the 40-60 ms pair was estimated from the psychometric function for AdaptWC-Q.*

	15-35 ms (fixed)	15-35 ms (adapt)	40-60 ms (fixed)	40-60 ms (adapt)
AR	84.2 (12.7)	84.4 (21.8)	45.6 (15.6)	66.1 (23.7)
DYS	74.5 (16.7)	78.6 (25.7)	44.1 (11.4)	57.4 (18.5)

Table 5: *Z-scores for individual participants within the DYS group on each of eight speech tasks. Performance that is below 1.65 standard deviation from the mean for the AR group is indicated in bold. The codes for individuals who were 'below norm' for three out of the eight speech tasks are in bold while the codes for individuals who are 'within-norm' on all eight tasks are italicized.*

Case	FixedAC	AdaptAC	FixedWC	AdaptWC	AdaptID-Q	AdaptID-N	WiN	WiNiCS
D1	**-4.24**	**-3.06**	-0.28	**-4.22**	-1.28	0.86	-0.28	-0.68
D2	**-2.08**	-0.81	-0.28	**-2.44**	**-2.90**	-1.56	0.16	-0.68
D3	**-2.80**	0.51	-0.13	0.16	-1.28	**-4.39**	**-1.66**	1.12
D4	**-1.72**	-0.02	-0.88	-0.66	**-3.15**	-0.60	**-2.50**	-1.28
D5	**-2.44**	0.77	-1.03	-0.53	-0.54	**-4.08**	**-2.19**	-1.28
D6	**-5.68**	**-2.80**	-0.13	-0.53	0.03	-1.25	0.36	-0.68
D7	**-5.32**	0.24	-0.58	0.16	-0.96	-1.31	**-2.30**	1.12
D8	**-2.80**	**-2.01**	0.02	-0.39	-1.02	-0.61	2.54	-0.68
D9	**-1.72**	0.77	-1.48	0.16	0.20	1.43	0.40	1.12
D10	-1.36	-0.81	0.17	**-2.17**	0.07	-0.39	-1.53	3.64
D11	-0.64	**-2.40**	-0.58	0.43	-1.27	1.38	-1.62	-1.28
D12	-0.64	**-2.93**	0.32	-1.48	-1.08	-0.74	-0.83	-0.68
D13	0.50	0.90	0.77	-0.39	0.99	0.56	-1.56	**-1.88**
D14	1.16	**-2.01**	-0.13	-0.39	-1.26	-1.43	-1.13	1.12
D15	1.52	0.77	-0.58	0.30	-0.16	2.20	**-3.39**	-0.08
D16	1.52	1.17	-0.43	0.71	-0.47	-0.01	-0.56	-0.68
D17	1.88	0.51	-1.33	-0.25	-0.91	0.55	-0.33	2.32
Total below norm	9	6	0	3	2	2	5	1

Table 6: *Z-scores for individual participants within the AR group on each of eight speech tasks. Performance below 1.65 standard deviation from the mean for the AR group is indicated in bold. The codes for individuals who were 'below norm' for three out of the eight speech tasks are in bold while the codes for individuals who are within-norm for all eight tasks are italicized.*

Case	FixedAC	AdaptAC	FixedWC	AdaptWC	AdaptID-Q	AdaptID-N	WiN	WiNiCS
AR1	**-3.88**	**-3.72**	-1.18	-0.53	**-1.75**	0.27	-1.17	-0.68
AR2	**-1.72**	0.38	0.32	**-1.90**	-0.35	-0.69	1.25	-0.32
AR3	**-3.88**	0.51	-1.33	-0.12	0.02	0.05	-0.48	-0.68
AR4	-0.64	**-2.67**	1.83	-0.94	-0.10	**-1.66**	0.30	-0.68
AR5	-1.36	0.64	1.68	0.16	**-2.86**	**-1.66**	-1.37	0.22
AR6	-0.28	-0.68	0.47	0.98	0.36	**-4.39**	-0.84	1.12
AR7	0.08	-0.95	0.17	-1.62	-0.20	-0.82	**-2.90**	0.52
AR8	0.80	-0.29	-0.28	-0.25	**-3.54**	-0.33	0.01	**-2.48**
AR9	0.80	-1.48	-0.28	0.84	**-1.77**	1.69	.	-0.08
AR10	0.80	1.57	-1.48	-1.35	1.04	1.52	**-2.01**	-1.28
AR11	1.88	-0.68	0.17	1.53	1.14	-0.23	-**2.02**	1.12
AR12	-1.36	0.90	0.17	0.43	0.95	-0.43	0.30	1.12
AR13	-0.64	0.38	-0.58	0.30	1.10	-1.22	-1.53	-0.68
AR14	-0.64	0.77	-0.28	-1.07	1.13	1.44	1.33	-0.68
AR15	-0.64	-0.68	1.08	-0.66	-1.35	0.30	0.27	0.22
AR16	0.44	0.24	0.17	-1.35	-0.55	0.58	1.34	1.12
AR17	1.16	0.77	-0.88	0.16	-0.06	0.47	1.31	-1.28
AR18	1.16	0.51	-1.63	0.57	-1.10	-0.22	-1.10	2.32
AR19	-0.28	-0.02	0.62	1.39	0.47	1.20	0.30	-0.08
AR20	0.44	0.77	1.23	1.53	1.03	-0.30	0.09	-1.28
Total below norm	3	2	0	1	4	3	3	1

Table 7: Individual scores for adults classified as ‘poor performers’ on the basis of being ‘below norm’ on at least three out of the eight speech tests. Scores are given for verbal and non-verbal IQ, phonological short-term memory and composite z-scores for reading, phonology, speech perception in quiet and in noise. Means are also given for the DYS ‘poor performer’ group, DYS ‘good performer’ group and for the AR group minus the one AR participant classified as ‘poor performer ‘. The composite z-scores for the AR ‘other’ group differ from a mean of 0 (s.d. 1). This is due to the fact that outliers were removed before the calculation of the AR mean used in the computing of individual z-scores for AR and DYS participants, but they are included in the mean scores below.

Case	Group	Verbal IQ	Non-verbal IQ	Phonological STM	READING	PHONOLOGY	QUIET	NOISY
AR1	AR	135	95	90	-0.73	-1.26	-2.47	-0.92
D1	DYS	135	110	95	-2.05	-0.45	-3.20	-0.06
D2	DYS	130	115	63	-3.74	-2.95	-2.06	-0.86
D3	DYS	114	145	65	-3.51	-3.00	-0.85	-1.46
D4	DYS	126	110	80	-2.88	-1.28	-1.39	-1.28
D5	DYS	102	115	75	-2.10	-2.98	-0.68	-2.10
DYS 'poor performers'	n=5	121.4 (13.3)	119.0 (14.7)	75.5 (13.0)	-2.9 (0.8)	-2.1 (1.2)	-1.6 (1.0)	-1.2 (0.8)
DYS 'other'	n=12	131.8 (22.0)	115.0 (16.9)	86.9 (12.9)	-2.7 (1.1)	-1.0 (1.4)	-0.6 (1.0)	-0.1 (0.6)
AR ‘other’	n=19	146.0 (17.1)	113.7 (12.1)	93.9 (5.5)	-0.1 (0.9)	-0.1 (0.9)	-0.1 (0.7)	-0.2 (0.7)