Exploring dyslexics' phonological deficit III: Foreign speech perception and production

Efstathia Soroli[1*], Gayaneh Szenkovits[2] & Franck Ramus[2]

[1] *Laboratoire Structures Formelles du Langage (UMR 7023, CNRS, Université Paris 8), Paris, France*

[2] *Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS, CNRS, DEC/ENS), Paris, France*

Abstract

The present study investigates French dyslexic and control adult participants' ability to perceive and produce two different non-native contrasts (one segmental and one prosodic), across several conditions varying short-term memory load. For this purpose, we selected Korean plosive voicing (whose categories conflict with French ones) as the segmental contrast, and lexical stress as the prosodic contrast (French does not use contrastive lexical stress). We also used a French (native) segmental contrast as a control. Tasks were either auditory discrimination or repetition of CVCV nonsense words. Short-term memory load was varied by presenting the stimuli either in isolation, in sequences of two, or in sequences of three. Our results show overall few differences between dyslexic and control participants. In particular, dyslexic participants performed similarly to controls in all tasks involving Korean plosives, whether in discrimination or in production, and regardless of short-term memory load. However, some group differences emerged with respect to lexical stress, in the discrimination task at greater short-term memory load. Various analyses suggest that dyslexic participants' difficulties are due to the meta-phonological nature of the task and to short-term memory load.

Keywords

Developmental dyslexia, second language acquisition, speech perception, lexical stress

* Correspondence should be addressed to Efstathia Soroli, Laboratoire Structures Formelles du Langage UMR 7023, CNRS, Université Paris 8, 59 rue Pouchet, 75017 Paris, France. E-mail: eva.soroli@sfl.cnrs.fr, telephone: +33(0)1.40.25.10.39, fax: +33(0)1.40.25.10.41.

Introduction

Developmental dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent written word recognition and by poor spelling and decoding abilities. These difficulties are often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction (Lyon, Shaywitz, & Shaywitz, 2003). Despite many theoretical debates, there is now a wide agreement that, at least for a majority of dyslexic children, the cognitive deficit underlying dyslexia lies mainly in the phonological domain, that is, in the ability to represent and process speech (Démonet, Taylor, & Chaix, 2004; Ramus, 2003; Snowling, 2000).

More specifically, more than thirty years of investigation of the "phonological deficit" have highlighted three broad areas of difficulties for dyslexic children: 1) phonological awareness, the ability to pay attention to, and consciously manipulate the units of speech (in particular the smallest ones: phonemes); 2) verbal short-term memory, the ability to retain phonological representations for a few seconds; 3) lexical retrieval, as tapped in rapid automatic naming tasks, where participants must retrieve the phonological forms of pictures (or colors, digits, or letters) in quick succession to name them as fast as possible (Wagner & Torgesen, 1987). Nevertheless the underlying nature of the phonological deficit is not well understood yet (Ramus, 2001).

Most theories of the phonological deficit share the assumption that dyslexic individuals' phonological representations are degraded in some way. Depending on the theory, they may be under- or poorly specified (Elbro, 1998; Snowling, 2000), more noisy (Harm & Seidenberg, 1999), have poorer temporal resolution (Tallal, Miller, & Fitch, 1993), or they might be insufficiently tuned to the native phonemic categories (Adlard & Hazan, 1998; Mody, Studdert-Kennedy, & Brady, 1997; Serniclaes, Sprenger-Charolles, Carré, & Démonet, 2001). Alternatively, it has also been proposed that dyslexic people's phonological representations may be intact, but more difficult to store or access under certain conditions (Ahissar, 2007; Ramus & Szenkovits, 2008).

In addition to these three main categories of symptoms, it is often reported that dyslexic pupils have inordinate difficulties learning foreign languages at school[1] (Downey, Snyder, & Hill, 2000; Helland & Kaasa, 2005; Service, 1992). Given that second language (L2) learning requires perceiving, paying attention to, memorising, and producing new speech sounds, dyslexic children's difficulties with phonological awareness and verbal short-term memory may provide a straightforward explanation for their difficulties with foreign languages. However, L2 learning is difficult for all learners, essentially as a function of the relationship between the learner's native linguistic system and the target system. This is particularly evident in the phonological domain, giving rise to foreign accent and to "language-specific listening" (Goto, 1971; Pallier, Christophe, & Mehler, 1997). Thus, beyond phonological awareness and verbal short-term memory problems, depending on the precise nature of their phonological deficit, dyslexic individuals may have more specific difficulties with respect to L2 acquisition.

Interestingly, different theories of the phonological deficit may make different predictions about L2 learning. Therefore, beyond the practical interest of understanding exactly what hinders dyslexic children's learning of foreign languages, this area of research may also have a genuine theoretical interest by shedding some light on the nature of their phonological deficit. In this study, we will therefore investigate the perception and production of foreign speech sounds by dyslexic and control students, in order to assess which theory makes the most accurate predictions. We now attempt to draw predictions with respect to L2 learning from a number of established or hypothetical theories of the phonological deficit.

Degraded phonological representations: If the hypothesis is that phonological representations are generally poorly specified (Adlard & Hazan, 1998; Elbro, 1996; Harm & Seidenberg, 1999; Mody et al., 1997; Snowling, 2000; Tallal et al., 1993), then presumably the poor specification should apply to foreign as well as to native speech sounds, leading to increased deficits in foreign speech sound categorisation and discrimination (relative to controls' already poor performance). Furthermore, according to most

[1] In this paper we are concerned only with late second language learning. Early (bilingual) language learning is a different problem, for which we are not aware that dyslexic children might have specific difficulties.

proponents of this view, some phonetic features are more at risk of being poorly represented than others: those that rely on fine acoustic distinctions, for instance stop consonants, as opposed to vowels or prosodic contrasts. Therefore, dyslexics' deficits should be relatively more marked for those contrasts in foreign speech. According to yet another version of this hypothesis, the degradation applies specifically to output phonological representations, which predicts that foreign speech production should be particularly affected (Hulme & Snowling, 1992).

Universal/allophonic phonological system: Under this hypothesis, would-be dyslexic infants fail to properly acquire the phonological categories of their native language. This would give rise to less sharp native categorical boundaries, as evidenced in identification and discrimination tasks. Thus, the dyslexic infant's phonological system would be less affected by exposure to a native language, and would therefore be closer to the universal state it presents at birth (Aslin, Werker, & Morgan, 2002; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Serniclaes, Van Heghe, Mousty, Carré, & Sprenger-Charolles, 2004). Consistent with this idea, Serniclaes and colleagues (2001; 2004) have suggested that dyslexics also retain a heightened sensitivity to universal phonetic boundaries not present in the native language, that are normally lost in normal phonological acquisition. This hypothesis makes interesting predictions with respect to L2 acquisition. Indeed, it predicts that, in cases where L2 phoneme categories coincide with universal boundaries that are not used in L1, dyslexics would have an advantage in categorising and discriminating sounds across these boundaries, relative to controls.

Beat perception deficit: According to Goswami and collaborators (2002), the phonological deficit is based on a deficit to perceive amplitude modulations of the acoustic signal that enhance speech segmentation in a way that is useful for reading acquisition. Contrary to most other hypotheses, the deficit is therefore thought to be related to the perception of a prosodic dimension rather than of fine phonetic properties. This hypothesis might therefore predict that difficulties in foreign speech learning would primarily affect prosodic rather than segmental contrasts.

Phonological access deficit: In a recent review, Ramus and Szenkovits (2008) have challenged the idea that dyslexic individuals' phonological representations are degraded at all, emphasising that deficits only appear as a function of certain task demands, e.g. metacognitive access, short-term memory load, or speeded access. They would therefore predict that dyslexics would have similar difficulties with foreign speech sounds as controls, but that their performance would become poorer only when those particular task demands increase.

It should be emphasised that, apart from the allophonic hypothesis (Serniclaes et al., 2004), none of the authors cited above have actually made explicit predictions about non-native speech processing. The above predictions therefore reflect what seems to us to follow most naturally from these hypotheses, based on our understanding of the descriptions provided. Proponents of these theories might of course disagree with our predictions and are welcome to further specify their own predictions in future work.

In the present study, we test dyslexic and control adult participants' ability to perceive and produce two different non-native contrasts (one segmental and one prosodic). For this purpose, we selected Korean stop consonant voicing (whose categories conflict with French ones) as the segmental contrast, and lexical stress as the prosodic contrast (French does not use contrastive lexical stress). We also used a French (native) segmental contrast as a control. Tasks were either auditory discrimination (tapping perception only) or repetition (tapping both perception and production). Finally, the material always consisted of CVCV nonsense words, which were presented either in isolation, in sequences of two, or in sequences of three, thereby varying short-term memory load.

Methods

Participants

Fifteen dyslexic and fifteen control adult participants with similar academic background and non-verbal IQ took part in this study. Participants were selected among those already recruited by Szenkovits and Ramus (2005). They were recruited through adverts in Parisian universities and received 10 € per hour of participation. Inclusion criteria were (1) to be a native, monolingual speaker of French aged above 18 years old, (2) to report no known neurological/psychiatric disorders or hearing impairment, and (3) to have a non-verbal IQ above 90. For controls, the crucial criteria was (4a) to report no known history of reading/oral language difficulties, and to have a reading age above the ceiling (14 years old) of our standardized reading test. For dyslexics, (4b) self- or institutional identification as a dyslexic person, and a reading score below the level of 14;6 years old (Grade 9).

All participants were given a questionnaire on their language background. They all reported French as their single native language, and they had all been exposed to French since birth. They were all late bilinguals (due to compulsory foreign language teaching in French schools), but none had learnt a second language before age 10, and none had lived in a foreign country for more than six months.

Procedure

Participants underwent a diagnostic battery during the first session to ensure that they met inclusion criteria, then the experimental tests in a second, separate session. All computerized tests and experiments were programmed, presented and scored on a personal computer using E-Prime (Schneider, Eschman, & Zuccolotto, 2002) for perception and DMDX (Forster & Forster, 2003) for production experiments. All experiments were carried out in a soundproof room. Stimuli were presented through headphones and responses were made on a response box.

Diagnostic battery

The diagnostic battery included intelligence and reading tests for the purpose of inclusion criteria. In addition, since we specifically targeted the phonological deficit to the exclusion of any other possible cause of dyslexia (e.g., purely visual), it included a set of classic phonological tasks and verified that all dyslexics had poor performance on those. Table 1 shows a summary of their demographic characteristics and of their performance on the diagnostic battery.

Insert Table 1 about here

Nonverbal intelligence was assessed by using Raven's Advanced Progressive Matrices Set I and Set II (Raven, Raven, & Court, 1998) in time-limited condition (40 minutes). Set I was used to familiarize participants with the test, Set II to calculate non-verbal IQ scores derived from the percentiles of United States norms (1993).

Receptive vocabulary was assessed with the EVIP test (Dunn, Thériault, & Dunn, 1993), a French Canadian version of the Peabody Picture Vocabulary Test – Revised.

Reading skills were assessed by the standardized French reading test "L'alouette" (Lefavrais, 1967). The text comprises 265 words ranging from common to rarely used words. Participants are instructed to read the text as fast and as accurately as possible. Standardized reading fluency scores are computed by combining total reading time and reading errors.

Orthographic skills were assessed with a speeded forced-choice task. Participants were presented successively with 24 triplets of words on computer screen. Each triplet included a correctly spelled word, and two misspelled versions. Participants had to press as quickly as possible the key corresponding to the correct spelling. Scores are the number of correct trials per second.

Digit Span: Forward and backward digit spans (from the French version of WAIS-III, Wechsler, 2000) were used to compute age-appropriate scaled scores, to obtain a measure of phonological working memory.

Spoonerisms: Participant were auditorily presented with pairs of words and were instructed to swap the first sound of each word, then pronounce the resulting pseudo-words while maintaining their correct order. A composite score taking into account both accuracy and speed is computed.

Rapid Automatic Naming: Participants completed three versions: picture and digit naming (2 sheets of 50 objects or digits) adapted from the Phonological Assessment Battery (Frederickson, Frith, & Reason, 1997), and color naming (2 sheets of 50 colors). Each naming test was administered twice with different sheets. The score is the sum of total naming time for both sheets. A composite RAN z-score was obtained by averaging z-scores from the three RAN tests.

In order to obtain a more synthetic view of participants' literacy and phonology skills, we additionally computed a composite literacy z-score from the mean z-scores of reading fluency and orthographic choice, and a composite phonology z-score from the mean z-scores of digit span, spoonerisms and RAN composite. Figure 1 shows the distribution of all participants on the composite measures of literacy and phonological tests.

One dyslexic and two control participants were excluded because they didn't meet all inclusion criteria. Performance in the diagnostic tests of the 27 remaining participants is reported in Table 1. One-Way ANOVAs show significant differences for all variables (vocabulary: $F(1,25)=7.86$, $p=0.01$; reading fluency: $F(1,25)=35.02$, $p<0.001$; orthographic choice: $F(1,25)=27.93$, $p<0.001$; literacy composite: $F(1,25)=41.23$, $p<0.001$; digit span: $F(1,25)=24.37$, $p<0.001$; spoonerisms: $F(1,25)=14.09$, $p=0.001$; RAN-objects: 23.43, $p<0.001$; RAN-digits: $F(1,25)=42.4$, $p<0.001$; RAN-colours: $F(1,25)=7.67$, $p=0.01$; RAN-composite: $F(1,25)=42.05$, $p<0.001$ and phonology composite: $F(1,25)=37.56$, $p<0.001$), apart from age and nonverbal IQ (both: $F(1,25)<1$) . Besides meeting the inclusion criteria based on reading fluency, all dyslexic participants scored at least 2.6 standard deviations below the control mean on the composite literacy z-score, and at least 1 SD below the control mean on the composite phonology z-score, thereby showing that they all had difficulties with phonological skills (see Figure 1).

Figure 1 about here

Experiment 1: Discrimination of native segments

For this experiment we used stimuli from Dupoux et al. (2001). These were 2 minimal pairs of CVCV pseudo-words [kupi]/[kuti] and [mipa]/[mita] recorded multiple times by two speakers (one male, one female). Stimuli were digitised at 16 kHz and 16 bits, digitally edited and stored on a computer disk. They were used to construct 32 sequences of 2 pseudo-words and 32 sequences of 3 pseudo-words.

In a first block of 16 trials, we used an AX discrimination task. Participants heard a sequence of 2 pseudo-words, followed by a 385 ms unintelligible babble noise, then a second sequence of 2 pseudo-words, that could differ (or not) from the first one by just one consonant (therefore by one phonetic feature: place of articulation) on either pseudo-word. Half the sequence pairs were the same, half were different. An example of a different trial is [kupi-kupi # kupi-kuti] where # refers to the babble noise. The babble noise was made of several superimposed speech sound tracks. This was to prevent participants from relying on echoic memory and to force them to encode the stimuli at the phonological, rather than acoustic, representation level. Across the two sequences, different recordings of a given pseudo-word were used, in order to maximise acoustic variability and therefore prevent discrimination on the basis of low-level acoustic cues.

In a second block of 16 trials, the task and the design remained the same but sequences were made of 3 pseudo-words, e.g., [mipa-mita-mipa # mipa-mita-mita].

Participants were asked to compare the two sequences and press a red key if the sequences were identical or a black key if they were different (the response-key mapping always remained on screen to prevent any confusion). The task started with 6 training trials with feedback, to ensure that participants understood the task and the nature of the contrasts to be discriminated, then followed with the two blocks of 16 trials. On average, the experiment lasted between 10 and 15 min.

This experiment was implemented as a control for all the other tasks involving foreign contrasts. In this experiment we didn’t feel the need to start with a block testing the discrimination of single pseudo-words, given that the native contrast was so easy for French listeners so that ceiling performance was expected even with sequences of 2 pseudo-words.

Experiment 2: Discrimination of Korean plosives

As the non-native segmental contrast, we selected the voicing contrast on Korean bilabial plosives. Unlike French (and English as well), which have two voicing categories for bilabial plosives (voiced [b] and unvoiced [p]), Korean has 3 categories: tense **pX** (/p͈/] in IPA), plain **p** (/p/ in IPA) and aspirated **ph** (/p^h/ in IPA). Instances of each of those 3 categories are typically perceived as [p]s of different acoustic qualities by French listeners, and are therefore easily confused (Ventureyra, Pallier, & Yoo, 2004). Informal listening suggests that pX and p are very difficult to distinguish, while ph, although perceived again as an instance of French p, is relatively easier to identify due to the strong aspiration cues. This is also confirmed by visual examination of the waveforms. We therefore predict that p-ph pairs will be more easily discriminated than pX-p pairs, as previously found by Ventureyra et al. (2004).

The material for this experiment was taken from the study by Ventureyra et al. (2004). It consisted in five triplets of Korean CVCV pseudo-words minimally differing in the voicing category of their initial consonant, e.g., [pXeda, peda, pheda] (see Appendix for the full list). Recordings were made by 3 male and 2 female Korean talkers of the Seoul dialect, in a sound-proof booth, low-pass filtered at 20 Khz and resampled at 16 bits/16 Khz. The mean duration of pseudo-words was 644 ms (SD=78 ms).

We used these pseudo-words to create sequences of one, two, or three pseudo-words (length factor). For lengths 2 and 3, pseudo-words from just one triplet ([pXeda, peda, pheda]) were used, and concatenated with a 0 ms inter-stimulus interval (ISI). The sex of speakers alternated within each sequence.

Sequences were presented in pairs, in a same/different discrimination task, with no time constraint for responses. At length 1, the two pseudo-words were different recordings by the same speaker and were played with a stimulus onset asynchrony (SOA) of 1000 ms. There were 36 trials, half same and half different, presented in a fixed pseudo-random order. At lengths 2 and 3, the two sequences were separated by a 400 ms babble noise. The speaker alternation was reversed between the two sequences, to further hinder the reliance on low-level acoustic cues. ‘Same’ trials included two sequences that were phonologically identical, but since each pseudo-word in the sequence was uttered by different speakers of opposite sex, they were acoustically different. ‘Different’ trials further differed by exactly one phonetic feature in one of the pseudo-words. At each length there were 16 trials, half same and half different, presented in a fixed pseudo-random order.

Prior to the experiment, participants were explained the existence of three categories of [p] in Korean., and were familiarized with them. Length 1 trials started with a short tutorial based on one triplet that was not used in the experiment. The 3 pseudo-words were printed on the screen and played simultaneously in the headphones, one after the other. Participants were then presented with the three written forms and had the possibility to replay each of them by pressing the corresponding key. They were allowed to listen to each exemplar five times before beginning the test phase. Length 2 trials were preceded with a short training of 6 trials with feedback to familiarise participants with the task. Length 3 trials were preceded by a warning that sequence length was about to increase to three pseudo-words. On average, this experiment lasted between 15 and 20 min.

Experiment 3: Production of Korean plosives

A subset of the disyllabic pseudo-words from Experiment 2 was used for this repetition task, and was presented either in isolation (length 1) or in pairs (length 2). At length 1, 9 different pseudo-words (3 for each category) were selected and played twice each. At length 2, 18 pairs of pseudo-words starting with a different [p] category were used. The experiment started with a short training using three pseudo-words not used in the test phase. Encouragement was provided regardless of performance (which could not be judged by the experimenter).

Participants were asked to repeat each pseudo-word or pair of pseudo-words as accurately as possible and were recorded on hard disk using a microphone. At length 2 they were informed that the two pseudo-words started with a different [p]. On average, the experiment lasted between 10 and 15 min.

All the pseudo-words recorded by participants were then excised and used in a subsequent experiment. This experiment was designed to have participants' productions judged by two Korean native speakers (both from Seoul). Each trial consisted of the presentation of the model Korean pseudo-word (or pair), followed by the repetition of one participant. Judges had to decide whether [p] sounds were correctly repeated or not by pressing an appropriate key. For each trial, each participant therefore received a score of 0 (incorrect), 1 (correct) or 0.5 (disagreement between judges). These scores were then averaged to produce an overall percentage of correct repetitions.

Experiment 4: Discrimination of lexical stress

Stress is used to distinguish different lexical items in many languages (such as Spanish or Greek), but not in French. Previous experiments have shown that French listeners have difficulties discriminating such contrasts, particularly when short-term memory load increases (Dupoux, Pallier, Sebastian, & Mehler, 1997; Dupoux et al., 2001).

For this experiment we used again stimuli from Dupoux et al. (2001). These were 2 minimal pairs of CVCV pseudo-words [kutì - kùti] and [mìpa – mipà] recorded multiple times by two speakers (one male, one female). Here pseudo-words of a given pair did not differ in terms of segmental content but in terms of the syllable that was stressed. Acoustic measurements of the stimuli indicated that stressed vowels differed from unstressed vowels in terms of duration, pitch and intensity (Dupoux et al., 2001). Stimuli were digitised at 16 kHz and 16 bits, digitally edited and stored on a computer disk. They were used to construct 16 sequences of 1, 2 and 3 pseudo-words respectively, concatenated with a 100 ms ISI.

Sequences were then presented in pairs in a same/different task, with no time constraint for responses. At length 1, the two pseudo-words were played with an SOA of 1000 ms. At lengths 2 and 3, the two sequences were separated by a 400 ms babble noise. ‘Same’ trials included two sequences that were identical from both segmental and prosodic points of view, but made with different recordings. ‘Different’ trials further differed by the location of stress in one of the pseudo-words. At each length, there were 16 trials, half same and half different, presented in a fixed pseudo-random order.

Length 1 trials started with a short tutorial on lexical stress based on three pairs of pseudo-words stressed on different syllables. Each pair was written on screen (with an accent marking stress) and played one after the other. Participants were then presented with the three written pairs and had the possibility to play each of them by pressing the corresponding key. They were allowed to hear each exemplar five times before going on to the test phase. Length 2 trials were preceded with a short training of 6 trials with feedback to familiarise participants with the task. Length 3 trials were preceded by a warning that sequence length was about to increase to three pseudo-words.

Experiment 5: Production of lexical stress

A subset of the disyllabic pseudo-words from Experiment 4 was used for this repetition task, and was presented either in isolation (length 1) or in pairs (length 2). At length 1, 3 different pseudo-words were selected and played with either stress position, using three different recordings, thus yielding 18 trials. At length 2, 18 pairs of the same pseudo-words were used. The experiment started with a short training using three pseudo-words. Encouragement was provided regardless of performance.

Participants were asked to repeat each pseudo-word or pair of pseudo-words as accurately as possible and were recorded on hard disk using a microphone. On average, the experiment lasted between 10 and 15 min. Participants’ recordings were then judged off-line by a native speaker of Greek. Stress cues were found to be sufficiently obvious not to require a second rater.

Results

For discrimination tasks, all percentages were converted into signal detection measures A' (sensitivity) and B''_D (bias) (Donaldson, 1992; Snodgrass & Corwin, 1988), based on hit rates (% correct detections of a difference) and on false alarm rates (% incorrect detections of a difference)[2]. A' scores are reported below for each discrimination experiment.

B''_D scores were computed for the 11 appropriate conditions and for each group. A significant group difference emerged in only one condition (the discrimination of native segments at length 2), showing a significant liberal bias for controls, but not for dyslexics. Given that this result does not seem interpretable, we assume that it arose by chance, so B''_D scores are not further analysed.

Performances in production tasks are reported as percentages of productions that were judged to be correct.

Experiment 1: Discrimination of native segments

Mean A' scores for each group at each length are reported on Table 2. We carried out a repeated-measures ANOVA with group (control, dyslexic) as between-subject factor and length (2 or 3 pseudo-words) as within-subject factor. The analysis showed a main effect of length ($F(1,25)=4.32$, $p=0.048$), but no effect of group ($F(1,25)=1.18$, $p=0.29$). Furthermore there was no group x length interaction ($F(1, 25)<1$).

Thus, sequences of two pseudo-words were better discriminated than sequences of three pseudo-words, and this did not differ between the two groups. This was expected given that this segmental contrast is extremely easy for French native speakers, and that the task did not put dyslexic participants at a specific disadvantage, neither in terms of short-term memory (sequences limited to 3 simple pseudo-words), nor in terms of acoustic features. These results therefore show that dyslexic participants are as able to perform

[2] If $H >= FA$, $A' = ½ + [(H – FA) (1 + H – FA)] / [4H (1-FA)]$. If $H <= FA$, $A' = ½ - [(FA – H) (1 + FA – H)] / [4FA (1-H)]$. A' varies between 0 and 1, with 0.5 indicating chance performance.

this simple task as control participants, at least when the phonological contrast to be discriminated is very familiar.

Insert Table 2 about here

Experiment 2: Discrimination of Korean plosives

Mean A' scores for each pair of plosives, for each group, and at each length are reported on Table 3. Mean A' scores for each group at each length, averaged across the two pairs are also reported on Table 2 for a more synthetic view. We carried out a repeated-measures ANOVA with group (control, dyslexic) as between-subject factor, pair (pX-p or p-ph) and length (1, 2 or 3 pseudo-words) as within-subject factors. The analysis showed a main effect of length ($F(2, 25)=30.7$, $p<0.001$), a main effect of pair ($F(2,25)=5.88$, $p=0.023$), but no effect of group ($F(1,25)<1$), and no significant interaction (all $F<2$, $p>.15$).

Thus, the p-ph pair was better discriminated than the pX-p pair. This confirms our expectations and is consistent with the results of Ventureyra et al. (2004). Interestingly, this pattern did not differ between the two groups. Furthermore, shorter sequences were overall better discriminated than longer ones. However, a closer examination of the data at each length suggests that performance is better at length 3 than at length 2 (see Table 2 and Figure 2). This difference is not statistically significant however ($t(26)=1.7$, $p=0.10$). This effect, if real, might reflect longer training at greater lengths, since conditions were always run from the smaller to the greater length. However a similar training effect is not observed across the same lengths in Experiment 4. It might equally be a random quirk in the data. At any rate, a similar profile is found for both groups, as evident in Figure 2, and therefore this does not endanger our assessment of group differences.

Insert Table 3 about here.

Experiment 3: Production of Korean plosives

Percent correct productions of each Korean p at length 1 are presented in Table 3. Such a detailed analysis has not been carried out at length 2 given that there were two target phonemes per trial, but percentages correct averaged across the three phonemes at lengths 1 and 2 are summarised on Table 2.

At length 1, we carried out a repeated-measures ANOVA with group (control, dyslexic) as between-subject factor, and phoneme (pX, p or ph) as within-subject factor. We found a main effect of phoneme ($F(2,25)=28.3$, $p<0.001$), but no effect of group ($F(1,25)<1$) and no group x phoneme interaction ($F(2,25)<1$). Paired t-tests revealed that ph was produced better than both pX and p (both $t(26)>5$, $p<0.001$), with no significant difference between p and pX ($t(26)<1$).

We also carried out another repeated-measures ANOVA collapsing across the three phonemes, with group (control, dyslexic) as between-subject factor, and length (1, or 2 pseudo-words) as within-subject factors. We found a main effect of length ($F(1, 25)=156$, $p<0.001$), but no effect of group ($F(1,25)<1$) and no group x length interaction ($F(1,25)=1.76$, $p=.2$).

These results therefore show that Korean ph is better produced than both p and pX by native French speakers, as expected from the distinctive acoustic features of ph. Furthermore, single pseudo-words are produced more easily than pairs of pseudo-words. This pattern did not differ between dyslexic and control participants.

Experiment 4: Discrimination of lexical stress

Mean A' scores for each group at each length are reported on Table 2. We carried out a repeated-measures ANOVA with group (control, dyslexic) as between-subject factor and length (1, 2 or 3 pseudo-words) as within-subject factor. The analysis showed a main effect of length ($F(2, 25)=15.1$, $p<0.001$), a main effect of group ($F(1,25)=6.42$, $p=0.018$), and no group x length interaction ($F(2, 25)=1.5$, $p=0.24$). Here, therefore, the results suggest that, besides the familiar length effect, dyslexic participants performed more poorly on average than control participants.

Experiment 5: Production of lexical stress

Percent correct productions of stressed pseudo-words at each length are reported in Table 2. We carried out a repeated-measures ANOVA with group (control, dyslexic) as between-subject factor and length (1 or 2 pseudo-words) as within-subject factor. The analysis showed a main effect of length ($F(2, 25)=4.82$, $p=0.038$), a marginally significant effect of group ($F(1,25)=3.98$ $p=0.057$), and no group x length interaction ($F(1, 25)<1$). Thus there is a trend in the same direction as for the discrimination of lexical stress, suggesting that French dyslexic individuals might have particular difficulties with lexical stress.

Global analysis

Figure 2 shows a summary of the performance of the two groups across all the conditions. It can be seen that the group difference varies across conditions, being statistically significant in two conditions, but with trends in the same direction in many other conditions. It therefore seems crucial to try and understand the factors that underlie such variations, and pose specific difficulties to dyslexic individuals.

Insert Figure 2 about here.

The present series of experiments varied, more or less systematically, a number of factors that may affect overall performance, or that may affect the performance of dyslexic relative to control participants. These are: the familiarity of the language (native or non-native), the nature of the contrast (segmental or suprasegmental), the short-term memory load (length varying from one to three pseudo-words), and the modality of the task (perception vs. production).

In order to try and disentangle which of these factors most affect the performance of dyslexic individuals relative to controls, we built a general linear model with performance scores as a dependent variable, each of the previously mentioned factors, plus group, as independent variables, and participant as a random variable. We modelled only main effects and the interaction between group and each of the other factors, in order to understand specifically which factors affect the group difference.

The analysis revealed main effects of all the factors: language (F(1, 287)= 244, p<0.001), contrast (F(1,287)=256, p<0.001), length (F(2,287)=45.9, p<0.001), modality (F(1,287)=12.5, p<0.001) and group (F(1, 71)=4.4, p=0.04). On the other hand, only one of the interactions tested was significant: group x contrast (F(1, 287)=7.64, p=0.006), all the other F values <1. Figure 3 illustrates the four interactions.

Insert Figure 3 about here.

Regarding the main effects, this analysis shows that performance was significantly higher for native than for foreign language, for suprasegmental than for segmental contrasts, for shorter than for longer sequences[3], and for discrimination than for production tasks. Furthermore, dyslexic participants performed significantly more poorly overall than controls.

Most interestingly, the analysis of interactions shows that dyslexic individuals were not more affected than controls by foreign vs. native speech, by sequence length (up to 3 pseudo-words), and in production vs. perception. However, they were significantly more affected by the contrast factor, that is, they performed relatively more poorly than controls for suprasegmental contrasts, compared to segmental contrasts.

Finally, in order to better understand which cognitive skills best predict dyslexic participants' difficulties with suprasegmental contrasts, we performed a multiple linear regression with, as dependent variable, lexical stress discrimination at length 3 (the condition showing the greatest group difference), and as regressors, non-verbal IQ, vocabulary, and the variables representing the main dimensions of the phonological deficit: spoonerisms (phonological awareness), digit span (verbal short-term memory), and the RAN composite score (lexical retrieval). When all regressors were entered simultaneously, only spoonerisms predicted a significant amount of variance[4]. The simple correlation with the dependent

[3] Notwithstanding the non-significant increase from length 2 to 3, already discussed p. 14.

[4] Interestingly, the next variable that almost succeeded entering the model (p=0.056) in a stepwise analysis was nonverbal IQ, not an additional phonological variable.

variable was 0.63 and the partial correlation 0.51. Spoonerisms alone predicted 40% of the variance of lexical stress discrimination at length 3. This result is not a trivial consequence of group differences, given that both RAN and digit span showed larger group differences than spoonerisms. This suggests that dyslexic participants' difficulties with this task are primarily explained by their phonological awareness deficit.

General discussion

In this study we investigated the ability of French adult dyslexic and control participants to discriminate and repeat non-native phonological contrasts. The contrasts investigated included Korean bilabial plosives, whose boundaries conflict with those of French bilabial plosives, and lexical stress, which is not used contrastively in French. Furthermore, we manipulated short-term memory load by varying the length of the sequences of pseudo-words to be discriminated or repeated. Our results show overall very few differences between dyslexic and control participants. In particular, dyslexic participants performed similarly to controls in all tasks involving Korean plosives, whether in discrimination or in production, and regardless of short-term memory load. On the other hand, some group differences emerged with respect to lexical stress, but only in the discrimination task and at greater short-term memory load.

To what extent can this pattern of results be attributed to ceiling and floor effects? In the easiest conditions (native contrast, length 1 for the stress contrast), there are undoubtedly ceiling effects (see Figure 2). In the most difficult conditions involving Korean plosive discrimination at lengths 2 and 3, performance is close to floor. Nevertheless, at length 1 for Korean plosives, performance is neither at floor nor at ceiling, and the two groups are identical. Furthermore, in the production tasks, whose chance level is very low, there is ample scope for group differences but none is observed. Therefore the conclusion that dyslexic participants do not differ from controls in their perception and production of Korean plosives cannot be due to floor or ceiling effects. The picture seems different for the stress contrast. Indeed, we see group differences only in the conditions where the short-term memory load draws performances below ceiling. In all the easier conditions, there is a trend for a group difference but it is likely that this is not statistically significant because of ceiling effects. Therefore, the conclusion that dyslexic participants have difficulties with the stress contrast holds despite ceiling effects, indeed it would be even stronger in the absence of ceiling effects in several conditions. To summarise, there are floor and ceiling effects in a few conditions, but they do not affect our main result, that dyslexic participants have difficulties with the stress contrast but not with Korean plosives.

Given this uneven profile of normal and poorer performance in dyslexic participants, it is of great theoretical interest to try and understand the factors that diminish their performance specifically in certain conditions. This was the point of the general linear model analysis that we carried out.

- Are dyslexic participants specifically impaired for non-native phonological contrasts, as opposed to native contrasts? No, in our analysis, the language factor did not interact with group.
- Are dyslexic participants more impaired in perception or in production? Again, our results suggest that the small group difference observed is identical across perception and production tasks.
- Can dyslexic participants' poorer performance be entirely explained by short-term memory load? No, the group difference did not systematically depend on this factor. However, it should be acknowledged that we did not push short-term memory abilities very far, with sequences of a maximum of 3 pseudo-words. Given the well-known difficulties of dyslexic individuals with verbal short-term memory (exemplified here in our digit span measure), it is of course expected that, with a sufficient short-term memory load, group differences would ultimately appear in all the tasks that we have used. Nevertheless, the fact remains that poor short-term memory is not sufficient by itself to explain the profile of performance observed in the present study.
- Finally, did dyslexic participants show poorer performance with segmental vs. suprasegmental contrasts? Indeed, this was the only factor that interacted with group, although perhaps in an unexpected direction. We found that dyslexic individuals had relatively greater difficulties with suprasegmental than with segmental contrasts.

Conceivably, these results may contribute to teasing apart or refine different theories of the phonological deficit in dyslexia. Most notably, theories that posit that dyslexic individuals' phonological representations are somewhat degraded, in particular at the finest temporal or spectral grain of representation (Adlard & Hazan, 1998; Elbro, 1996; Harm & Seidenberg, 1999; Mody et al., 1997;

Snowling, 2000; Tallal et al., 1993), would have predicted the poorest performance in tasks involving fine acoustic distinctions, i.e., tasks involving the Korean segmental contrasts. It is not clear at all how they could predict our present pattern of results, i.e., normal performance on Korean plosives but relatively poorer performance on a suprasegmental contrast that is instantiated by massive intensity, duration, and pitch cues spreading over hundreds of milliseconds.

The particular version of this theory held by Hulme and Snowling (1992), according to which *output* phonological representations are specifically degraded, should probably have predicted specific difficulties in production tasks. Although our production tasks did not specifically tap output representations (as they involve input representations as well, and the link between input and output), there is no doubt that they did engage output representations, more than our discrimination tasks did. But our results do not suggest that this involvement of output phonological representations created a particular difficulty for dyslexic participants.

Concerning the allophonic perception hypothesis (Hoonhorst et al., 2009; Serniclaes et al., 2001; Serniclaes et al., 2004), our results unfortunately do not allow for a clear-cut conclusion. This hypothesis has so far been developed only for the voicing continuum, as measured by VOT, and predicts that dyslexic individuals may discriminate better speech sounds that span a universal voicing boundary. Here, the Korean plosives that we used did differ in terms of voicing, and did span at least one of the universal boundaries (that at +30 ms VOT). However, it seems that Korean plosives may also differ by other acoustic cues than just VOT (Abramson & Lisker, 1972), so we cannot absolutely certify that dyslexic and control participants' similar performance has been achieved by exploiting exactly the same acoustic cues. This hypothesis therefore deserves further, more specifically-designed investigations. However it does not straightforwardly predict our pattern of results.

Interestingly, the beat perception theory (Goswami, 2006; Goswami et al., 2002) could have predicted a specific deficit with suprasegmental contrasts. Indeed this theory posits that dyslexic children have difficulties perceiving the amplitude rise that signals syllable onsets. And amplitude (or intensity) is one of the cues indicating stress (at least in our material). Nevertheless it is not entirely clear if a deficit in the detection of amplitude rise time would predict the pattern of results that we have obtained. Indeed, a deficit in the ability to precisely detect amplitude rise time is not the same thing as a deficit in generally perceiving amplitude over the duration of an entire syllable. In our material, stressed and unstressed vowels differed by 1.6 dB on average. Can the deficit hypothesised by Goswami et al. hinder the perception of such differences? The question remains open. Furthermore, our stimuli included other cues to stress (duration and pitch), which were potentially usable even by a participant who would be entirely unable to perceive amplitude cues. Therefore, although the beat perception theory superficially seems to predict our pattern of results, it is not clear that it really does, or at least the theory would need to be worked out in greater details in order to make specific predictions about the ability to discriminate stress contrasts.

On the other hand, there is additional, if sparse, evidence that dyslexic individuals may have difficulties perceiving some acoustic cues to prosody, such as frequency and amplitude modulations (Witton, Stein, Stoodley, Rosner, & Talcott, 2002), as well as some difficulties with speech rhythm perception and production (Wolff, 2002; Wood & Terrell, 1998) and lexical stress (de Bree, Wijnen, & Zonneveld, 2006; Wood, 2006). Thus, although there is not a so-to-speak "prosodic theory" of the phonological deficit in dyslexia, there is certainly some evidence that the phonological deficit manifests in the prosodic domain, among others.

According to yet another hypothesis, dyslexic people's phonological representations are intact, but access to these representations is limited under various task constraints, particularly those involving explicit awareness, short-term memory or rapid retrieval (Ramus & Szenkovits, 2008). This hypothesis therefore predicts that, across the present set of experiments, dyslexic participants' relative performance should simply vary as a function of these task constraints. One of these constraints, verbal short-term memory, was specifically manipulated by varying sequence length. However, the general linear model analysis revealed that variations in sequence length did not explain our pattern of results.

Thus it seems that none of the standard theories of the phonological deficit, at least as currently formulated in the literature, can immediately explain our results. In order to understand our results, a finer analysis of both the material and the tasks we used seems necessary.

Korean plosives are phonemic categories that are in conflict with French phonemic categories. This is known to be the most difficult situation for a non-native listener (Best, 1994; Best, McRoberts, & Goodell, 2001), one for which there is no good solution short of abandoning one's native language. Lexical stress is very different. Although French does not use stress to differentiate lexical items, it does have stress at the end of words, and it does modulate it across the sentence, to mark phrase-final words, or to produce focus for example. Thus French listeners have great difficulties perceiving and producing stress at a different position than the last syllable of the word, nevertheless their perceptual system has certainly not become insensitive to the acoustic cues of stress. Their problem is more to realise that they need to use those cues contrastively to discriminate and produce different words, and to automatise this process. The different nature of the problems posed by Korean plosives vs. lexical stress is illustrated in our data by the higher overall performance for the latter than for the former contrast (main effect of contrast in the GLM, see Fig. 3a).

With Korean plosives, dyslexic and control participants faced equally the nearly impossible task to try to ignore one's native phonemic categories and categorise sounds according to a conflicting boundary. Their performance was equally poor, getting close to floor in discrimination as soon as length increased to 2 pseudo-words. On the other hand, with the stress contrast, dyslexic participants may have found it more difficult to reflect on the acoustic cues supporting lexical stress, and to attend to them in order to perform the task. This may have been particularly taxing with the addition of a second difficulty factor, verbal short-term memory. In other words, we are suggesting that there is an important meta-phonological component to the task of dealing with lexical stress for French listeners, and that the poorer performance of dyslexic participants in those tasks may be explained by the combination of their poorer phonological awareness and their poorer verbal short-term memory. This conjecture is supported by our multiple regression analysis of performance in stress discrimination at length 3, showing as unique predictor spoonerisms, a primarily phonological awareness task with a working memory component.

Finally, what are the consequences of our results regarding foreign language learning by dyslexic individuals? Acquiring non-native sounds is undoubtedly a difficult task for everybody, with difficulty varying as a function of the relationship between native and non-native sound categories (Best, 1994). However, our results do not support the idea that dyslexic individuals have greater difficulties with non-native sounds than controls do. They certainly do have inordinate difficulties with late second language acquisition (Downey et al., 2000; Helland & Kaasa, 2005), but this does not seem to be explained by specific difficulties with non-native sounds. Rather, this seems to be better explained by their deficits in phonological awareness and verbal short-term memory, two cognitive skills that are highly recruited in second language learning (Service, 1992).

Acknowledgements

We wish to thank Christophe Pallier and Sharon Peperkamp for kindly providing their stimuli for this study, Emmanuel Dupoux, Willy Serniclaes and LSCP members for much discussion and feedback. We are especially grateful to Sofiane Gueddana and Nikoletta Soroli for help in experiment preparation and running. This work was supported by the Fyssen Foundation.

References

Abramson, A. S., & Lisker, L. (1972). *Voice timing in Korean stops.* Paper presented at the 7th International Congress of Phonetic Sciences, Montréal

Adlard, A., & Hazan, V. (1998). Speech perception in children with specific reading difficulties (dyslexia). *Quarterly Journal of Experimental Psychology, 51A*(1), 153-177.

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends Cogn Sci, 11*(11), 458-465.

Aslin, R. N., Werker, J. F., & Morgan, J. L. (2002). Innate phonetic boundaries revisited. *J Acoust Soc Am, 112*(4), 1257-1260; author reply 1261-1254.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech* *perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *J Acoust Soc Am, 109*(2), 775-794.

de Bree, E., Wijnen, F., & Zonneveld, W. (2006). Word stress production in three-year-old children at risk of dyslexia. *Journal of Research in Reading, 29*(3), 304-317.

Démonet, J. F., Taylor, M. J., & Chaix, Y. (2004). Developmental dyslexia. *Lancet, 363*(9419), 1451-1460.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121*(3), 275-277.

Downey, D. M., Snyder, L. E., & Hill, B. (2000). College students with dyslexia: persistent linguistic deficits and foreign language learning. *Dyslexia, 6*(2), 101-111.

Dunn, L. M., Thériault, C. M., & Dunn, L. M. (1993). *Échelle de vocabulaire en images Peabody (EVIP)*. Circle Pines: American Guidance Service.

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language, 36*, 406-421.

Dupoux, E., Peperkamp, S., & Sebastian-Galles, N. (2001). A robust method to study stress "deafness". *Journal of the Acoustical Society of America, 110*(3), 1606-1618.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing: An Interdisciplinary Journal, 8*, 453-485.

Elbro, C. (1998). When reading is "readn" or somthn. Distinctness of phonological representations of lexical items in normal and disabled readers. *Scand J Psychol, 39*(3), 149-153.

Forster, K. L., & Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav Res Methods Instrum Comput, 35*(1), 116-124.

Frederickson, N., Frith, U., & Reason, R. (1997). *Phonological Assessment Battery*. Windsor: NFER-NELSON.

Goswami, U. (2006). Sensorimotor impairments in dyslexia: getting the beat. *Developmental Science, 9*(3), 257-259.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., et al. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proc Natl Acad Sci U S A, 99*(16), 10911-10916.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia, 9*, 317-323.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychol.Rev., 106*(3), 491-528.

Helland, T., & Kaasa, R. (2005). Dyslexia in English as a second language. *Dyslexia, 11*(1), 41-60.

Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2009). French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology, 104*(4), 353-366.

Hulme, C., & Snowling, M. (1992). Deficits in Output Phonology - an Explanation of Reading Failure? *Cognitive Neuropsychology, 9*(1), 47-72.

Lefavrais, P. (1967). *Test de l'Alouette* (2ème ed.). Paris: Editions du Centre de Psychologie Appliquée.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1-14.

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: auditory processing or phonological coding? *J Exp Child Psychol, 64*(2), 199-231.

Pallier, C., Christophe, A., & Mehler, J. (1997). Language-specific listening. *Trends in Cognitive Sciences, 1*(4), 129-132.

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

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*(2), 212-218.

Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology, 61*(1), 129-141.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide.* Pittsburgh: Psychology Software Tools Inc.

Serniclaes, W., Sprenger-Charolles, L., Carré, R., & Démonet, J. F. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language and Hearing Research, 44*, 384-399.

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

Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology, 45A*(1), 21-50.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34-50.

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

Szenkovits, G., & Ramus, F. (2005). Exploring dyslexics' phonological deficit I: Lexical vs. sub-lexical and input vs. output processes. *Dyslexia, 11*(4), 253-268.

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Ann.N.Y.Acad.Sci., 682*, 27-47.

Ventureyra, V. A. G., Pallier, C., & Yoo, H. Y. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics, 17*, 79-91.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192-212.

Wechsler, D. (2000). *WAIS-III: Echelle de l'intelligence de Wechsler pour adultes - Troisième Edition*. Paris: Les Editions du Centre de Psychologie Appliquée.

Witton, C., Stein, J. F., Stoodley, C. J., Rosner, B. S., & Talcott, J. B. (2002). Separate Influences of Acoustic AM and FM Sensitivity on the Phonological Decoding Skills of Impaired and Normal Readers. *J Cogn Neurosci, 14*(6), 866-874.

Wolff, P. H. (2002). Timing precision and rhythm in developmental dyslexia. *Reading and Writing, 15*(1-2), 179-206.

Wood, C. (2006). Metrical stress sensitivity in young children and its relationship to phonological awareness and reading. *Journal of Research in Reading, 29*(3), 270-287.

Wood, C., & Terrell, C. (1998). Poor readers' ability to detect speech rhythm and perceive rapid speech. *British Journal of Developmental Psychology, 16*, 397-413.

Tables

Table 1. Results of the diagnostic battery.

Group		Age	Non verbal IQ[a]	Vocabulary[b]	Reading fluency[c]	Orthographic choice[d]	Digit span[e]	Spoonerisms[f]	RAN[g]				Composite Scores	
									Objects	Digits	Colours	Composite	Literacy[h]	Phonology[i]
		ns	*ns*	**	***	***	***	***	***	***	**	***	***	***
Controls	Mean	26.69	112.73	164.62	70.08	0.67	11.92	0.14	56.22	26.21	47.84	0	0	0
	SD	5.14	13.36	1.71	6.01	0.12	2.81	0.07	7.18	3.63	9.13	1	1	1
Dyslexics	Mean	26.5	111.61	158.92	132.79	0.38	7.36	0.06	74.93	38.62	59.21	-2.78	-10.5	-2.2
	SD	3.92	12.1	7.12	37.71	0.16	1.95	0.05	12.09	5.91	11.89	1.21	5.81	0.87

ns: not significant

** p<=0.01

*** p<=0.001

[a] Ravens' matrices, Standard Scores.

[b] EVIP raw scores.

[c] Adjusted reading time (s) for the French 'Alouette' reading test.

[d] Orthographic choice raw scores (words correct / s).

[e] WAIS-III FR Scaled scores.

[f] Percentage correct responses divided by average response time (s).

[g] Average naming time of the two sheets of each Rapid Automatic Naming test.

[h] ‘Literacy’ is the average of reading and orthography z-scores.

[i] ‘Phonology’ is the average z-score of all phonological tests: digit span, spoonerisms and rapid naming.

Table 2. Main results for each task at each length

Group		Native segment discrimination		Korean segment discrimination			Korean segment production		Stress discrimination			Stress production	
	Length	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**1**	**2**	**3**	**1**	**2**
Control	Mean	0.97	0.94	0.83	0.54	0.59	0.64	0.4	0.95	0.9	0.83	0.99	0.94
	SD	0.03	0.06	0.08	0.17	0.14	0.15	0.14	0.07	0.1	0.17	0.05	0.08
Dyslexic	Mean	0.94	0.9	0.8	0.54	0.63	0.64	0.36	0.89	0.79	0.68	0.91	0.84
	SD	0.08	0.14	0.09	0.21	0.18	0.06	0.13	0.08	0.15	0.22	0.16	0.19
Total	Mean	0.95	0.92	0.81	0.54	0.61	0.64	0.38	0.92	0.84	0.75	0.95	0.89
	SD	0.06	0.11	0.08	0.19	0.16	0.11	0.13	0.08	0.14	0.21	0.12	0.15

Scores are A’ values for discrimination tasks and % correct for production tasks.

Table 3. Discrimination of each pair of Korean plosives at each length. and production of each Korean plosive.

Group		Discrimination						Production		
		pX-p	p-ph	pX-p	p-ph	pX-p	p-ph	pX	p	ph
	Length	**1**	**1**	**2**	**2**	**3**	**3**	**1**	**1**	**1**
Control	Mean	0.83	0.82	0.5	0.56	0.48	0.68	0.53	0.45	0.92
	SD	0.1	0.07	0.19	0.25	0.29	0.15	0.31	0.17	0.11
Dyslexic	Mean	0.81	0.79	0.46	0.61	0.64	0.62	0.53	0.54	0.87
	SD	0.13	0.08	0.26	0.24	0.25	0.25	0.26	0.17	0.18
Total	Mean	0.82	0.81	0.48	0.59	0.56	0.65	0.53	0.5	0.89
	SD	0.11	0.07	0.23	0.24	0.28	0.21	0.28	0.18	0.15

Scores are A’ values for discrimination tasks and % correct for production tasks.

Figures

Figure 1. Scatter plot of composite phonology and literacy z-scores.

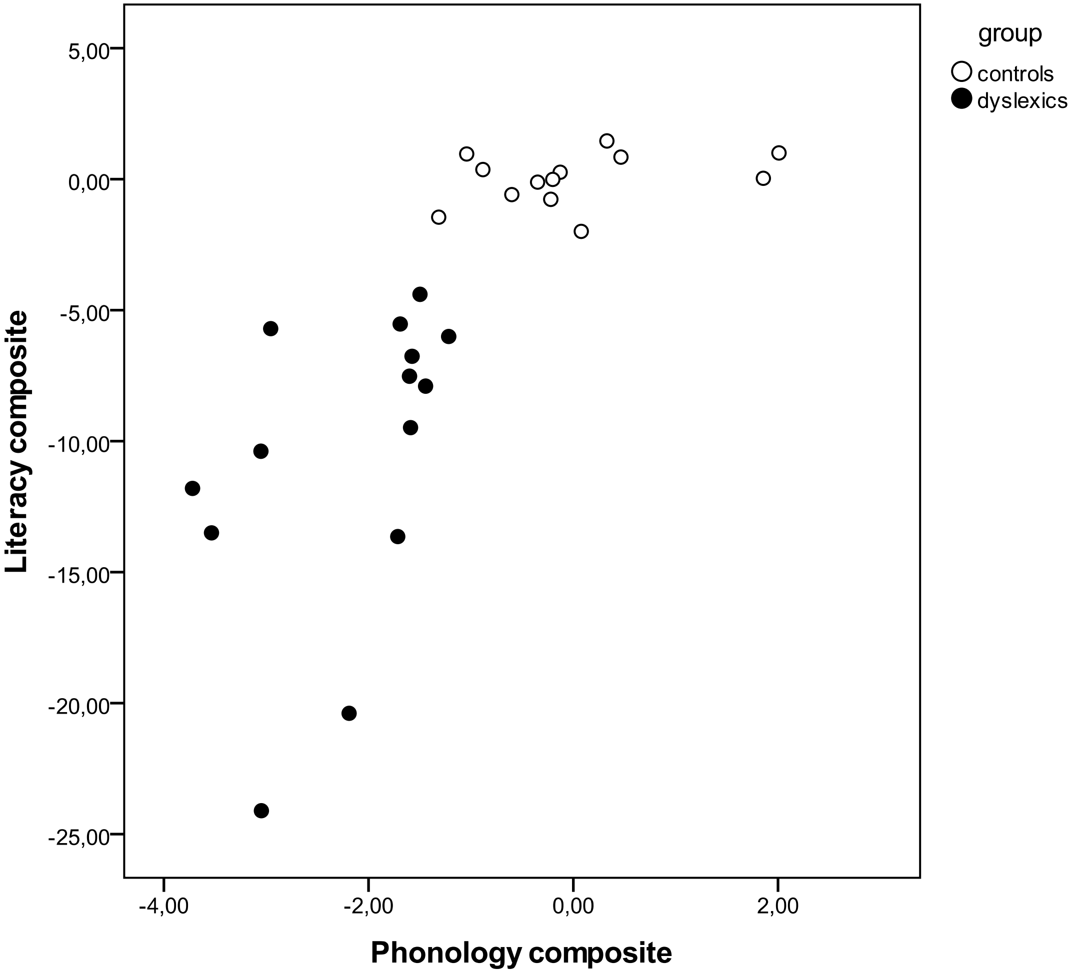


Figure 2. Summary performance in all conditions

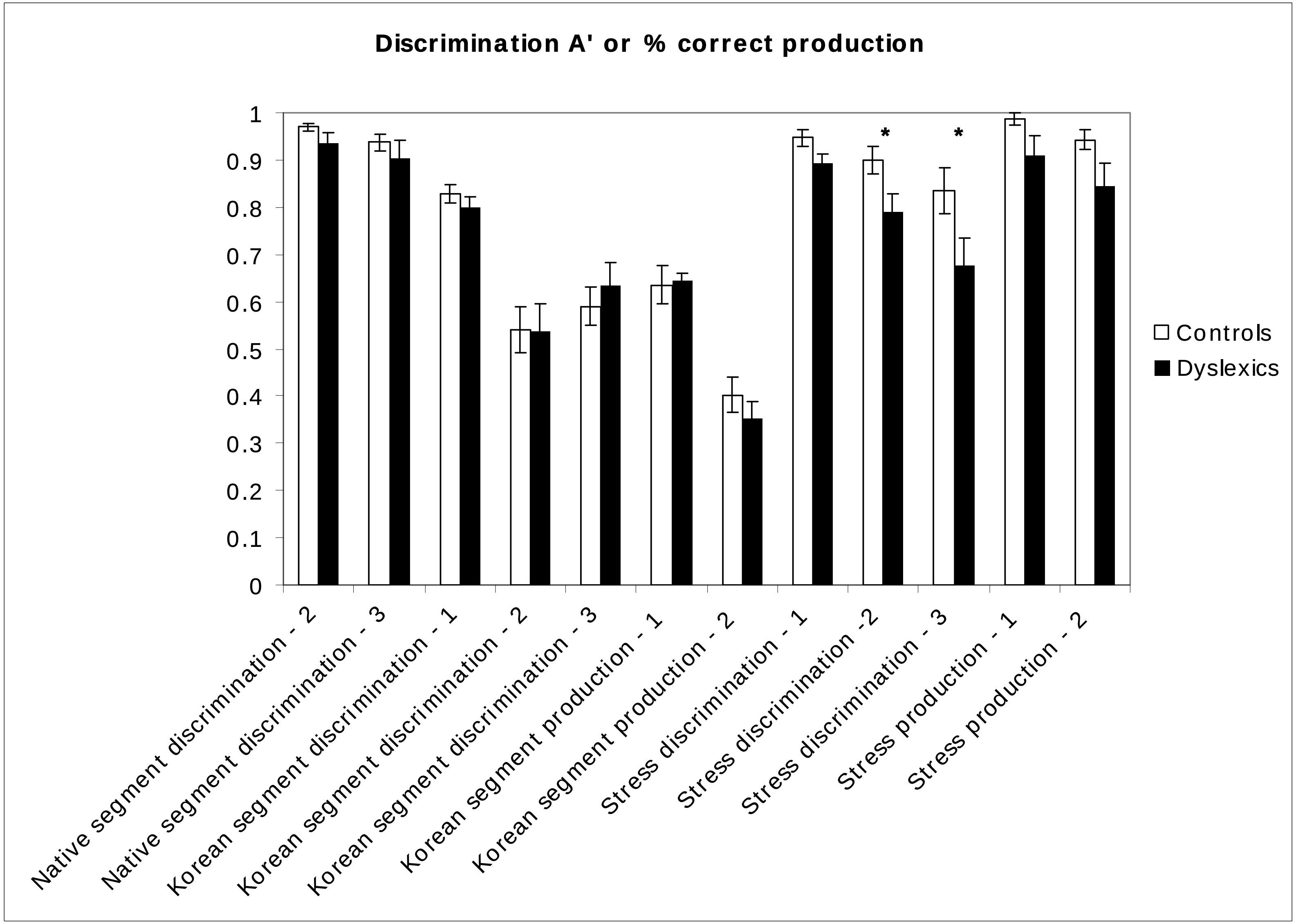
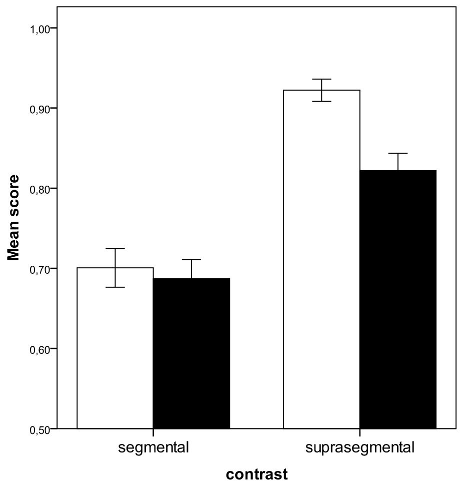
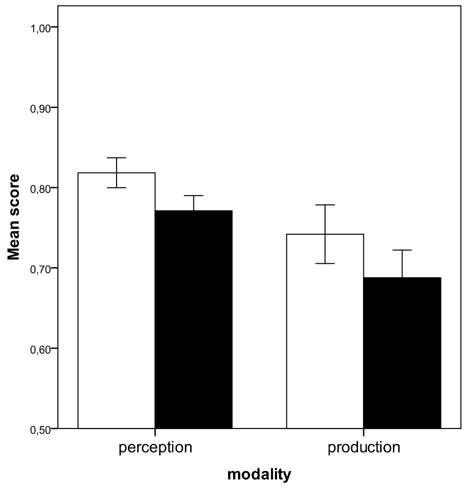


Figure 3. Interaction plots for the general linear model.

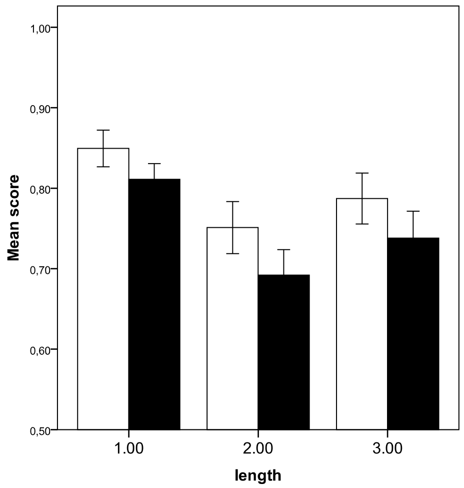
a)



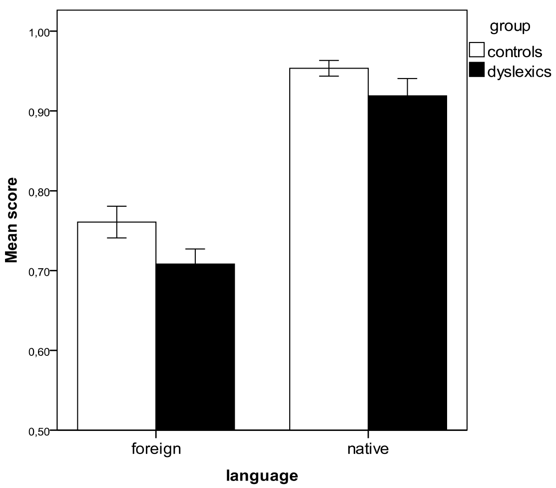
b)



c)



d)



Appendix

Material

Korean plosives

pXada	pada	phada
pXaga	paga	phaga
pXeda	peda	pheda
pXida	pida	phida
pXiga	piga	phiga
pXore	pore	phore
pXuga	puga	phuga
pXuri	puri	phuri

Native plosives

Mipa	mita
Kupi	kuti
Paku	patu

Lexical stress

Mìpa	mipà
Mìta	mità
Pàku	pakù
Pàtu	patù
Kùpi	kupì
kùti	kutì