Journal of Memory and Language 64 (2011) 59–73

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Rise time perception and detection of syllable stress in adults with developmental dyslexia

Victoria Leong, Jarmo Hämäläinen, Fruzsina Soltész, Usha Goswami *

Centre for Neuroscience in Education, University of Cambridge, United Kingdom

ARTICLE INFO

Article history:
Received 6 April 2010
revision received 17 August 2010

Keywords:
Dyslexia
Syllable stress
Rise time
Rhythm
Phonology

ABSTRACT

Introduction: The perception of syllable stress has not been widely studied in developmental dyslexia, despite strong evidence for auditory rhythmic perceptual difficulties. Here we investigate the hypothesis that perception of sound rise time is related to the perception of syllable stress in adults with developmental dyslexia.

Methods: A same-different stress perception task was devised and delivered to a sample of 40 adults in two formats, one using pairs of identical 4-syllable words and one using pairs of two different 4-syllable words. Auditory perception of rise time, frequency and intensity, and phonological awareness, phonological memory and reading were also measured.

Results: We show that adults with dyslexia performed significantly more poorly in both versions of the stress perception task. Individual differences in the perception of rise time were linked to the accuracy of performance.

Conclusions: To our knowledge this is the first direct demonstration of syllable stress perception deficits in dyslexia. The accurate perception of intonational patterning and rhythm may be critical for the development of the phonological lexicon and consequently for the development of literacy. Even high-functioning compensated adults with dyslexia show impairments in speech processing.

Introduction

Developmental dyslexia is a neurodevelopmental condition found across languages, for which the cognitive hallmark is impaired phonological processing (Snowling, 2000; Ziegler & Goswami, 2005). Evidence that this hallmark "phonological deficit" is related to impaired basic auditory processing has been accumulating during the last decade, in studies of both alphabetic and non-alphabetic languages. The auditory parameter most consistently found to be impaired has been perception of the amplitude envelope onset (rise time), or its correlate, amplitude modulation depth (Corriveau, Pasquini, & Goswami, 2007; Goswami, Fosker, et al., 2010; Goswami, Gerson, & Astruc, 2009; Goswami et al., 2002; Goswami, Wang, et al., 2010; Hämäläinen, Leppänen, Torppa, Muller, & Lyytinen, 2005; Hämäläinen, Salminen, & Leppänen, in press; Hämäläinen et al., 2009; Lorenzi, Dumont, & Fullgrabe, 2000; Muneaux, Ziegler, Truc, Thomson, & Goswami, 2004; Pasquini, Corriveau, & Goswami, 2007; Richardson, Thomson, Scott, & Goswami, 2004; Rocheron, Lorenzi, Fullgrabe, & Dumont, 2002; Suranyi et al., 2009; Thomson, Fryer, Maltby, & Goswami, 2006; Thomson & Goswami, 2008). Behaviourally, rise time is most closely associated with the perceptual experience of speech rhythm and stress (Hoequist, 1983; Morton, Marcus, & Frankish, 1976). However, to date, there has been no investigation of the possible relationship between basic auditory processing of rise time and the perception of syllable stress in spoken words in dyslexia. A clear prediction of the "rise time" theory of developmental dyslexia (Goswami et al., 2002) is that the perception of syllable stress should be impaired in individuals with dyslexia, and that individual differences in rise

---

* Corresponding author. Address: Centre for Neuroscience in Education, 184 Hills Rd., Cambridge CB2 8PQ, United Kingdom. Fax: +44 1223 767602.

E-mail address: ucg10@cam.ac.uk (U. Goswami).

0749-596X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2010.09.003

time perception should predict the severity of any impairment in perceiving syllable stress.

Despite the lack of direct evidence for a stress perception impairment in dyslexia, recent studies using reiterative speech tasks are consistent with the prediction that individuals with developmental dyslexia should be impaired in perceiving syllable stress. Regarding adults with developmental dyslexia, Kitzen (2001) developed a reiterative speech task in which each syllable in a word was converted into the same syllable (here DEE). This enabled distinctive phonetic information in words and phrases to be removed while retaining the stress and rhythm patterns of the originals. Kitzen converted film and story titles into ''DeeDees'', so that (for example) ''Casablanca'' became DEEdeeDEEdee (STRONG weak STRONG weak, or SWSW). Adolescent participants with dyslexia heard a tape-recorded DeeDee sequence while viewing three alternative (written) choices, for example ''Casablanca'', ''Omega Man'' and ''The Godfather''. Kitzen found that her participants with dyslexia were significantly poorer in this choice task than age-matched controls. She also reported that performance in the DeeDee measure was significantly associated with syllable and phoneme segmentation skills, and with word reading abilities and reading comprehension. In logistic regression analyses carried out to predict group membership (dyslexic versus control), the DeeDee measure was a highly significant predictor of group status (along with syllable segmentation and rapid naming measures). All three measures together predicted group membership with 97% accuracy (phoneme segmentation was not a significant predictor). However, one drawback with this study was the use of written response choices for participants who had difficulties in processing written language.

Goswami et al. (2009) developed two DeeDee tasks suitable for children with dyslexia, which avoided reading demands (see also Whalley & Hansen, 2006). In their tasks, children saw a picture of a ''famous name'' familiar to British participants (such as the English footballer David Beckham) or pictures corresponding to familiar film and book titles (such as Harry Potter). Familiarity with the pictures was assessed in a pretest. During the experimental sessions, the children were asked to select which of two ''DeeDee'' phrases that they heard matched the picture. For example, the correct match for ''Harry Potter'' was DEEdeeDEEdee (SWSW). Goswami et al. reported that the children with dyslexia (who were aged on average 12 years) performed significantly more poorly than age-matched controls in both the 'Film and Book Titles' and 'Famous Names' DeeDee tasks. Performance in the DeeDee tasks was also a significant predictor of reading development in the sample, for example individual differences in the 'Famous Names' task accounted for 25% of unique variance in reading accuracy after controlling for age and IQ. DeeDee perception predicted reading even when phonological awareness (performance in a rhyme oddity task) was additionally controlled (still accounting for 16% of unique variance, $p < .001$). Finally, individual differences in measures of the auditory perception of rise time predicted unique variance in the reiterative speech task.

One drawback of reiterative speech tasks is that they require participants to derive an abstract representation of the stress patterning of a particular utterance rather than to perceive the stress patterns in the utterance directly. Studies of direct stress perception in non-dyslexic adults have used a variety of experimental paradigms, including visual and auditory lexical decision, shadowing tasks, speech gating tasks, and word recognition of mis-stressed words (see Cutler (2005), for a recent review). As discussed by Cutler (2005), prior information about stress patterning does not seem to facilitate lexical access in English, although in some studies stress information helps to resolve lexical competition. For example, Cooper, Cutler, and Wales (2002) showed using a fragment priming task that information about syllable stress helped listeners to assign initial syllables to source words such as *admiral* versus *admiration*. The adults heard sentences like ''The speech therapist said.'' and then had to make a lexical decision about the target words (e.g., admiral/admiration). The auditory primes were fragments of complete words pronounced with either first syllable stress (''ADmir'') or third syllable stress (''admir'' from *admiration*). Cooper et al. reported that a fragment like ''ADmir'' activated *admiral* more than *admiration*, while a fragment like ''admir'' activated *admiration* more than *admiral*. Their conclusion was that English adults do make use of suprasegmental information in recognising spoken words. Slowiaczek (1990) asked participants to listen to spoken words that were mixed with white noise and were presented with either correct stress (e.g., SPECulative) or incorrect stress (specUlative). Participants had to write down what they heard and were credited for accurate word recognition. Slowiaczek found no effects of mis-stressing in this recognition task. However, when she asked participants to shadow what they heard in a subsequent experiment, there was an effect of mis-stressing on response *speed*. Participants were slower to produce the mis-stressed words, suggesting that lexical stress is coded as part of the phonological representation.

As the cognitive difficulty in developmental dyslexia lies in the accurate neural representation of the phonological information in words, stress perception may be expected to play an important role in the development of well-specified phonological representations. English is a free-stressed language, as prominence may occur on different syllables, falling at different positions in different words (as in ''orNATE'' for the isolated word versus ''ORnate BALcony'' for continuous speech). Studies of early phonological development in English suggest that infants and very young children adopt a primarily lexical strategy to stress placement, that is they learn stress as part of the phonological representation of a particular word (e.g., Klein, 1984). However, many English words used with infants and young children follow a strong–weak pattern (*mummy*, *daddy*, *baby*, *doggie*), and so it is possible that template learning plays a role in the development of knowledge about stress. In general, strong syllables are louder and longer than weak syllables, and have a higher pitch (frequency). Jusczyk, Houston, and Newsome (1999) reported that infants could segment words with strong–weak patterns by 7½ months of age, but appeared to mis-segment words following a weak–strong pattern. For example, if the infants heard a sentence such as ''*her guitar is too fancy*'', they segmented ''*taris*'' as a plausible word (treating "*taris*" rather than "*guitar*" as familiar during the dishabituation test). By 10½ months of age, infants did not make these mistakes. Sensitivity to the predominant stress patterns of English words is clearly important for segmenting words and syllables from the speech stream, and therefore for phonological representation (see also Echols, 1996; Mattys & Juscyk, 2001).

Recent theories of developmental phonology have also suggested an important role for prosodic sensitivity in explaining phonological development (Gerken, 1994; Pierrehumbert, 2003; Vihman & Croft, 2007). For example, Pierrehumbert (2003) argued for early-acquired "prosodic structures" as the basis for language acquisition, proposing a model based on the acquisition of complex language-specific exemplars from the input that were stored in rich phonetic and prosodic detail (see also Port, 2007). She argued that phonetic perception is dependent on the prosodic context. Indeed, stress perception studies with both children and adults have suggested that target phonemes are detected more efficiently when they are in stressed syllables (e.g., Mehta & Cutler, 1988; Wood & Terrell, 1998). Therefore, current evidence suggests that stress is an integral part of the phonological representations of English words developed by infants, and that phonological development is characterised by an inter-dependency of phonetic and prosodic information.

It thus seems plausible to propose that the phonological difficulties experienced by children and adults with developmental dyslexia must involve reduced sensitivity to stress and intonational patterning as well as reduced sensitivity to phonological units like syllables, onsets, rimes and phonemes. As noted, the auditory correlates of stress are most usually defined as involving amplitude, duration and frequency. Classical theories (e.g., Fry, 1954) accorded fundamental frequency the key role in stress perception, with duration and intensity (amplitude) playing secondary roles. More recent investigations using natural speech have shown that amplitude and duration cues play a stronger role in prosodic prominence than fundamental frequency (Choi, Hasegawa-Johnson, & Cole, 2005; Greenberg, 1999; Kochanski, Grabe, Coleman, & Rosner, 2005). For example, Greenberg (1999) described an automatic prosodic algorithm developed to label stressed and unstressed syllables in a corpus of spontaneous speech. The algorithm depended on three separate parameters of the acoustic signal, duration, amplitude and fundamental frequency. In contrast to classic accounts, Greenberg reported "fundamental frequency turns out to be relatively unimportant for distinguishing between the presence and absence of prosodic prominence… the results indicate that the product of amplitude and duration … yields the performance closest to … linguistic transcribers" (p. 172). Similar conclusions were reached by Kochanski et al. (2005) in an investigation of a large corpus of natural speech covering 7 English dialects.

Greenberg (2006) has explicitly linked changes in rise time to prosodic prominence by proposing a theory of how the "energy arc" of speech (the linguistic manifestation of the energy arc is the syllable) is produced by manner of articulation. By this account, the energy contour of the speech signal is an arc rising to a peak in the nucleus of each syllable and then descending. Rise time (the rate of change in intensity or signal energy as the nucleus of the syllable is produced by the articulators) should be particularly critical for stress perception. The specific way in which the arc ascends to the peak depends on whether the syllable is stressed (here more energy is produced) and the phonetic composition of the syllable onset – with more sonorous onsets, speakers take longer to reach the peak. Prosody thus affects both the height and length of the energy contour, and so the amplitude envelope of speech reflects the prosodic properties of speech.

Loudness (amplitude) perception per se is not usually impaired in studies of auditory processing in developmental dyslexia. Rather, perception of the *rate of onset* of changes in amplitude (rise time) is impaired. For example, the different cohorts of children with developmental dyslexia tested by Richardson et al. (2004), Thomson and Goswami (2008) and Goswami et al. (2009) did not exhibit significantly raised auditory thresholds for amplitude compared to age-matched controls in two forms of a two-interval forced choice (2IFC) task. In one version of this intensity threshold task, the children were asked to judge which of two sounds A and B was softer (Richardson et al., 2004). In the second version, the children heard two sequences of five sounds (AAAAA versus ABABA), and had to detect which sequence varied in intensity (Goswami et al., 2009; Thomson & Goswami, 2008). Group thresholds for intensity discrimination were statistically equivalent for children with dyslexia and age-matched controls in all three studies. Nevertheless, individual differences in the ABABA intensity discrimination task were predictive of performance in the "Film and Book Titles" reiterative speech task, an indirect measure of sensitivity to syllable stress, accounting for 18% of unique variance after controlling for age and IQ (Goswami et al., 2009). Similarly, in the Thomson and Goswami (2008) study, intensity discrimination was significantly correlated with performance in a Tempi discrimination task even when non-verbal IQ was controlled (the Tempi task asked children to judge which of two cartoon bears playing trumpets were producing notes at a slower pulse rate, $r = .39$, $p < .01$). Therefore, if outcome measures involve an element of periodicity, as in the DeeDee task and in Tempi detection, intensity discrimination may be a significant predictor of individual differences in addition to rise time. The relationship of intensity discrimination to perceiving syllable stress patterns in multi-syllabic words remains to be tested (although see Foxton, Riviere & Barone, 2010 for an audio-visual stress recognition task in which amplitude perception did play a role in detecting *visual* prosody).

These relationships between simple intensity discrimination and periodicity are consistent with a more recent study of developmental dyslexia using a musical metrical perception task based on simple tunes comprised of strong and weak "beats" (Huss, Verney, Fosker, Fegan, & Goswami, 2010). In this musical study, children with dyslexia aged 10 years and control children were asked to judge whether two short tunes were the same or different in metrical structure. The tunes varied in metrical complexity (e.g., a 6-note tune in duplex time with takt on the first note, versus a 15-note tune in 4/4 time with takt on the second note). Huss et al. found that the children with dyslexia were impaired in perceiving metrical similarity irrespective of the metrical complexity of the different tunes. The severity of the children's metrical perceptual difficulties was uniquely predicted by performance in only two of the basic auditory processing tasks that were administered, rise time discrimination and intensity discrimination. Pitch and duration thresholds did not predict unique variance in the metrical perception task in block-entry multiple regression equations, despite the fact that metrical dis-similarity depended on inserting longer durations between adjacent musical notes. As metrical structure is a focus of interest in linguistic studies of syllable stress, with metrical structure accorded an important organisational role in determining syllable, word and clausal boundaries, the difficulties of individuals with dyslexia in metrical perception are again consistent with the difficulty hypothesised here in perceiving syllable stress in dyslexia. In fact, metrical perception accounted for 42% of unique variance in reading in the musical metre study, making it a stronger predictor of reading development in this sample of children than phonological awareness.

Accordingly, we assume here that very basic auditory processes are used in perceiving metrical structure in both music and language, and that individual differences in these basic auditory processes affect individual differences in the extraction of periodic structure and accordingly the perception of syllable stress in speech. To test this hypothesis, we measured basic auditory processing in a sample of adults with and without developmental dyslexia, and we also measured stress perception in a same-different task based on 4-syllable words. From our analysis of over 2500 4-syllable words drawn from the CELEX database, we found that 4-syllable words in English most commonly receive primary stress on the second syllable. Forty-four per cent of words (like *maternity* and *ridiculous*) conform to this typical stress template, which can be denoted as '0200'. The remainder of words either received primary stress on the first syllable (24%), as in *difficulty* and *military* (2000 stress template), or on the third syllable (28%), as in *comprehensive* and *interaction*. These words also had secondary stress on the first syllable (1020 stress template).

In the current study, we used only words with first or second syllable stress, that is, 2000 or 0200 template words. We recorded a female British speaker saying tokens of each type of word (2000 or 0200 template) with either correctly or incorrectly placed stress. For example, two tokens of the word *maternity* were recorded, one with correctly placed stress as in *maTERnity* (WSWW) and the other with incorrectly placed stress as in *MAternity* (SWWW). We then paired these tokens in all four possible ways (SWWW–SWWW, WSWW–WSWW, SWWW–WSWW, WSWW–SWWW). Participants were asked to judge whether the two tokens in the pair contained the "same" or "different" stress patterns. We also varied whether the spoken token was the same word in each pair (as in *maternity–maternity*, Experiment 1, thereby keeping segmental phonology constant), or was two different words with matching syllable stress templates (as in *maternity–ridiculous*, Experiment 2, thereby conceptually more similar to the DeeDee task, in that abstract stress templates must be compared to make a judgement). We were interested to see whether participants with dyslexia would find it more difficult to make judgements about shared syllable stress in each experiment. Note that in both experiments, stress pattern similarity can nevertheless be judged "on-line" using the acoustic information in the heard tokens, without recourse to the mental lexicon.

Experiment 1

Method

Participants

Twenty adults with developmental dyslexia (11 male; mean age 25.3 years, range 17.5 years – 41.8 years) and twenty adults without dyslexia (7 male; mean age 26.3 years, range 18.1 years – 38.5 years) participated in the study. Eighteen of the adults with dyslexia had a formal statement of developmental dyslexia, the remaining two participants showed severe literacy and phonological deficits according to our own test battery which was administered to all participants. As phonological deficits were part of the inclusion criteria for the study, it is possible that participants whose difficulties were visual and not phonological were excluded from the sample. All participants had no diagnosed additional learning difficulties (e.g. dyspraxia, ADHD, autistic spectrum disorder, speech and language impairments) and spoke English as a first language. Participant details are shown in Table 1. All participants took part in both Experiments 1 and 2 on separate days, with Experiment 1 being performed first.

Tasks

Standardised ability tests. All participants were given 2 subscales of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), a non-verbal subscale (Block Design) and a verbal subscale (Vocabulary). Literacy skills were assessed using the untimed Wide Range Achievement Test (Reading and Spelling scales, WRAT-III, Wilkinson, 1993). A measure of short-term memory, the Weschler Adult Intelligence Scale-Revised forward digit span subtest was also administered (WAIS-R; Weschler, 1998).

Table 1
Participant details.

Group	Dyslexic	Controls	$F(1, 38)$
Chronological age (years)	25.3	26.3	.32
(*sd*)	(5.6)	(5.2)	
WRAT reading standard score	102.5	114.7	23.44***
(*sd*)	(10.0)	(5.3)	
WRAT spelling standard score	97.8	115.6	38.21***
(*sd*)	(11.3)	(6.2)	
WASI vocabulary subscale *T* score (mean = 50)	63.1	64.7	.64
(*sd*)	(6.3)	(5.9)	
WASI block design subscale *T* score (mean = 50)	59.0	61.0	.85
(*sd*)	(7.4)	(6.3)	
WAIS-R digit span subscale score (out of 16)	10.5	12.2	5.28*
(*sd*)	(2.6)	(2.0)	

* $p < .05$.
*** $p < .001$.

Phonological awareness measures.

i. *Spoonerisms.* This task was drawn from the Phonological Assessment Battery (PhAB; Fredrickson, Frith, & Reason, 1997). Participants heard 10 pairs of words presented orally by the experimenter. Participants were asked to swap the onset phonemes of the pair of words (e.g. for "*sad cat*"; subject responded "*cad sat*"). Scores on this measure were out of a possible 20 points.

ii. *RAN (Rapid Automatized Naming).* Two versions of an object RAN task designed originally for children were administered, one based on pictures of objects whose names resided in dense phonological neighbourhoods (RAN Dense: *Cat*, *Shell*, *Knob*, *Thumb*, *Zip*), and one based on pictures of objects whose names resided in sparse phonological neighbourhoods (RAN Sparse: *Web*, *Dog*, *Fish*, *Cup*, *Book*). Participants were shown a sheet of paper with the same pictures repeated 50 times. In each case, they were asked to produce the names as quickly and accurately as possible. Performance was timed, and the two tasks were combined to give an average RAN score in seconds.

Psychoacoustic tasks. The psychoacoustic stimuli were presented binaurally through headphones at 74 dB SPL. The auditory tasks were presented using an adaptive staircase procedure (Levitt, 1971) with a combined 2-up 1-down and 3-up 1-down procedure; after 2 reversals, the 2-up 1-down staircase procedure changes into 3-up 1-down. The step size halves after the 4th and 6th reversal. A test run typically terminates after 8 response reversals or alternatively after the maximum possible 40 trials. Four attention trials were randomly presented during each test run, using the maximum contrast of the respective stimuli in each auditory task. The threshold score achieved was calculated using the mean of the last four reversals.

i. *Amplitude Envelope Onset (Rise Time) Task (1 Rise).* This was a rise time discrimination task in AXB format. Three 800 ms tones were presented on each trial, with 500 ms ISIs. Two (standard) tones had a 15 ms linear rise time envelope, 735 ms steady state, and a 50 ms linear fall time. The third tone varied the linear onset rise time logarithmically with the longest rise time being 300 ms. Participants were introduced to three cartoon dinosaurs. It was explained that each dinosaur would make a sound and that the task was to decide which dinosaur's sound was different from the other two and had a softer rising sound (longer rise time, this was either sound A or B, never sound X). As an integral part of the software programme feedback was given after every trial on the accuracy of performance. Schematic depiction of the stimuli can be found in Richardson et al. (2004).

ii. *Frequency task*. This was a frequency discrimination task also delivered in an AXB format. The standard was a pure tone with a frequency of 500 Hz presented at 74 dB SPL, which had a duration of 200 ms. The maximum pitch difference between the stimuli presented in this task was 60 Hz. Participants were introduced to three cartoon elephants. It was explained that each elephant would make a sound and that the task was to decide which elephant's sound was higher.

iii. *Intensity task*. This was an intensity discrimination task delivered in a 2IFC format. The standard was a pure tone with a frequency of 500 Hz presented at 74 dB SPL, which had a duration of 200 ms. The intensity of the second tone ranged from 54 to 74 dB SPL. Participants were introduced to two cartoon mice. It was explained that each would make a sound, and the task was to decide which sound was softer. Participant's performance on phonological awareness and psychoacoustic tasks are shown in Table 2.

Syllable stress task. This task was based on 20 4-syllable words with lexical templates that had first syllable stress (2000, such as *caterpillar* and *difficulty*) and 20 4-syllable words with lexical templates that had second syllable stress (0200, such as *maternity* and *ridiculous*). The words were selected from an initial list of more than 2500 4-syllable words with first and second syllable-stress pooled from two linguistics databases (MRC Psycholinguistic Database and CELEX). The words were selected on the basis of syllable structure (no consonant clusters in the first two syllables), spoken and written frequency, and overall familiarity. Words also did not have alternative pronunciations. The full list of stimuli is presented as Appendix A. The words were divided into two lists of 20 words each (each list comprising 10 words with 2000 lexical templates and 10 words with 0200 lexical templates). Participants received one word list in Experiment 1 and the other in Experiment 2, which were given on separate days, with order of presentation of the word lists counterbalanced across participants. The two lists, and the two sets of lexical templates (2000, 0200), were matched as closely as possible for spoken and written frequencies. Mean values for 2000 templates were Cobuild spoken frequency 21.7

Table 2
Group performance in the phonological and auditory tasks.

Group	Dyslexic	Controls	*F*(1, 36)
Spoonerisms[a]	15.2	17.8	7.70[b,**]
(*sd*)	(3.2)	(2.3)	
RAN time in seconds	35.2	30.3	11.84[b,**]
(*sd*)	(4.5)	(4.0)	
Auditory threshold			
1 Rise in ms	63.0	40.3	11.50[**]
(*sd*)	(28.0)	(5.5)	
Frequency in Hz	12.5	9.1	5.20[*]
(*sd*)	(5.5)	(3.5)	
Intensity in dB	2.1	1.9	[c].80
(*sd*)	(0.9)	(0.4)	

[a] Score out of 20.
[b] Degrees of freedom are (1, 34).
[c] Degrees of freedom are (1, 32).
[*] $p < .05$.
[**] $p < .01$.

(*sd* 22) and written frequency 288.6 (*sd* 294.2). Mean values for 0200 templates were Cobuild spoken frequency 15.5 (*sd* 30) and written frequency 224.3 (*sd* 315.2). Neither difference was statistically significant, $F(1, 38)$ for spoken frequency = 0.44, $F(1, 38)$ for written frequency = 0.55.

All items were produced naturally by a native female speaker of British English and recorded for computerised presentation using Audacity and Praat software. Two spoken tokens were recorded for each word. In one token, the speaker emphasised only the first syllable of the word (producing a SWWW stress pattern). In the other token, the speaker emphasised only the second syllable of the word (producing a WSWW stress pattern). This resulted in a total of 80 spoken tokens from 40 words. Word pairs were then created for each trial by combining the two spoken tokens in all four possible ways. The recorded tokens were analysed for mean intensity, duration, amplitude rise time and F0. Mean values for unstressed or stressed first syllables (such as *ma* or *MA* in *maTERnity* and *MAternity* respectively) and stressed or unstressed second syllables (such as *TER* or *ter* in *maTERnity* and *MAternity* respectively) are shown in Table 3. The values shown confirm that the acoustic parameters differed consistently between stressed and unstressed syllables across different words on a paired samples *t*-test. On average, stressed syllables were higher in intensity and pitch, and had longer durations and slower rise times than unstressed syllables. These acoustic differences are illustrated in Fig. 1, which shows a 3D plot of the amplitude envelopes for the word pair *DIfficulty and diFFIculty*. To create the figure, sound stimuli were first bandpass filtered into 12 logarithmically-spaced channels spanning a frequency range from 100 to 4000 Hz. Each frequency channel was then demodulated individually to extract its amplitude envelope. The figure plots time on the *x*-axis, frequency on the *y*-axis, and amplitude on the *z*-axis. Marked with arrows on the plot are duration, onset rise and intensity changes for stressed and unstressed versions of the syllable '*ffi*'. Differences in the frequency profile (circled) are also apparent as the stressed '*FFI*' shows larger amplitudes in mid-frequency channels than the unstressed '*ffi*'.

Table 3
Acoustic parameters of stressed and unstressed syllables (mean across 40 words).

	Stressed	Unstressed	$t(39)$
First syllable manipulated	E.g. MA in *MAternity*	E.g. ma in *maTERnity*	
Median intensity in dB	73.2	71.2	4.89^{***}
(*sd*)	(5.1)	(4.2)	
Duration in ms	181.4	148.1	5.58^{***}
(*sd*)	(61.9)	(51.9)	
Amplitude rise time in ms	94.3	82.5	3.17^{**}
(*sd*)	(35.3)	(33.8)	
Mean F0 in Hz	243.5	209.2	9.77^{***}
(*sd*)	(23.3)	(15.6)	
Second syllable manipulated	E.g. TER in *maTERnity*	E.g. ter in *MAternity*	
Median intensity in dB	72.3	70.1	5.03^{***}
(*sd*)	(4.3)	(4.6)	
Duration in ms	175.3	145.2	5.14^{***}
(*sd*)	(58.9)	(50.3)	
Amplitude rise time in ms	95.5	79.2	3.10^{**}
(*sd*)	(43.3)	(37.4)	
Mean F0 in Hz	241.8	199.3	11.64^{***}
(*sd*)	(22.4)	(14.7)	

** $p < .01$.
*** $p < .001$.

During task presentation, participants simply heard a word pair where two word tokens were presented one after the other. Participants were told to make same-different judgments about the position of syllable stress in the pair (such as *MIlitary* – *miLItary* [different] or *MIlitary* – *MIlitary* [same]). There was a 500 ms interval between the words in a pair, and a 2000 ms interval between trials after a response was given. Participants responded by pressing right or left buttons on the keyboard. The side of the same/different buttons was randomised across participants. Participants were told to respond as quickly and accurately as possible after they saw a question mark appear on the screen, which appeared at the end of the second word. Reaction time was recorded as the time from the question mark appearing to the participant's response (correctly-answered trials only). Feedback on the correctness of the response was provided on each trial by showing either a 'happy' smiley cartoon icon (correct response), or a 'pirate' cartoon icon (incorrect response). Apart from the '?' prompt and feedback icons, the computer screen remained blank whenever auditory stimuli were being presented. Prior to starting the experiment, participants received four practice trials. Note that participants were instructed to judge whether the position of stress was on the same or different syllables, not whether the word tokens were correctly or incorrectly pronounced. Participants did not report any difficulty in understanding what judgement was required.

There were four possible types of word pairs which differed in stress position, SWWW–SWWW, WSWW–WSWW, both requiring "same" judgements, and SWWW–WSWW, WSWW–SWWW, both requiring "different" judgements. Examples of these pairs are given in Fig. 1.This factor is referred to as Same/Different Judgement. In Experiment 1, the words were either based on 10 tokens with first syllable stress lexical templates (e.g., *difficulty–difficulty*) or were based on 10 tokens with second syllable stress lexical templates (e.g., *maternity–maternity)*. This factor is referred to as First/Second stress template. Combining this factor with the four types of word pairs created 80 trials, which were fully randomised and presented in two 40-trial blocks. The experiment therefore used a 2 × 2 × 2 design (Group × First/Second × Same/Different Judgement). The experimental design is summarised in Fig. 2.

Results

Auditory discrimination and phonological awareness data were explored by group to check that assumptions of normality (skew and kurtosis) were met. The SPSS boxplot function was used to check for outliers, and any data points lying farther than three interquartile ranges from

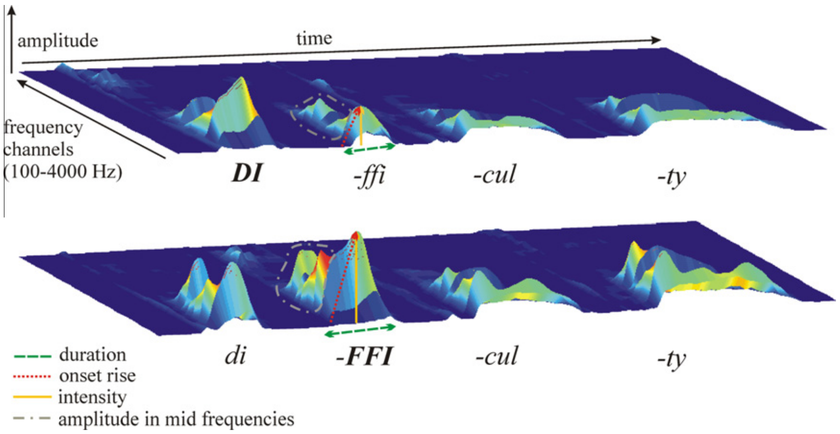


Fig. 1. Amplitude envelope across frequencies for the word *difficulty* produced with stress on the first or second syllable.

	Lexical template (First/Second)	Trial type (Same/Different)	Examples
1	First syllable stress template (2000)	SAME	*DIfficulty – DIfficulty* *diFFIculty – diFFIculty*
2		DIFFERENT	*DIfficulty – diFFIculty* *diFFIculty – DIfficulty*
3	Second syllable stress template (0200)	SAME	*maTERnity – maTERnity* *MAternity – MAternity*
4		DIFFERENT	*maTERnity – MAternity* *MAternity – maTERnity*

Fig. 2. Schematic depiction of design of Experiment 1.

the further edge of the box were removed. Five outlier scores were identified and removed for the auditory processing tasks (2 control scores for 1 Rise, 2 dyslexic scores for Frequency, 1 control score for intensity). Group data for the standardised tasks is provided in Tables 1 and 2. As would be expected given previous work, the participants with dyslexia were significantly less sensitive to auditory rise time and to frequency than their controls, but were not significantly different for intensity discrimination. Participants with dyslexia were significantly impaired in all the reading measures, and were also significantly impaired in the phonology measures. These differences were established using a series of one-way ANOVAs ($N = 40$), and F and p values are reported in Tables 1 and 2.

Mean performance (% correct and reaction time) for making judgements about shared syllable stress in each condition, as well as calculated d' and criterion values are shown in Table 4. Preliminary analyses confirmed that reaction times did not differ between groups and response time is not analysed further. Paired t-tests for d' and c values revealed significant group differences on both measures. Participants with dyslexia showed a significantly lower sensitivity (d') than controls ($t(1,38) = 2.7$, $p = .01$) on the task. They were also more biased toward giving a 'same' response than controls ($t(38) = -3.2$, $p = .004$). This indicates that participants with dyslexia had more difficulty detecting acoustic differences between two items that were stressed differently, sometimes mistaking them as having the same stress pattern.

In order to check the effects of varying the syllable template, a 2 × 2 ANOVA (Group × First/Second syllable stress) was carried out, taking d' as the dependent variable. As would be expected, this showed a significant main effect of Group, $F(1,38) = 7.3$, $p = .010$. However, the effect of

Table 4
Group performance on the stress perception task in Experiment 1: Mean% correct, Mean RT, d' and c (*sd* in parentheses).

	% Correct		RT in ms	
	Dyslexic	Control	Dyslexic	Control
First syllable stress template (2000)				
Same judgement	98 (3.0)	98 (4.1)	1085 (233)	1046 (292)
Different judgement	94.8 (8.7)	99 (2.1)	1069 (224)	1040 (312)
Second syllable stress template (0200)				
Same judgement	98.3 (3.4)	98.8 (2.2)	1068 (231)	9882 (286)
Different judgement	92.3 (7.9)	98.5 (2.9)	1044 (207)	1051 (309)
d' (sensitivity)	4.3 (0.6)	4.7 (0.3)		
Criterion (bias)	0.2 (0.3)	0.0 (0.1)		

First/Second stress template was not significant, $F(1, 38) = .81$, $p = .372$, and there was no interaction between First/Second syllable stress and Group, $F(1, 38) = 1.2$, $p = .281$. The results suggest that the participants with dyslexia found it difficult to judge shared stress when an identical item was pronounced with two different stress patterns, whether the stress template was SWWW or WSWW.

In order to examine whether these stress perception difficulties were related to inefficiencies in auditory perception, multiple regression analyses were used. Three 2-step fixed order equations were computed, all entering Group at Step 1 and then either rise time threshold, frequency threshold or intensity threshold at Step 2. The dependent variable in each case was d'. The results are shown in Table 5. As can be seen, rise time discrimination contributed 24% of unique variance to judgements about syllable stress. Frequency and intensity discrimination did not contribute significant unique variance to stress judgements, even though frequency discrimination also differed significantly between the two groups of participants. The data suggest a unique relationship between basic auditory perception of rise time and the accurate perception of syllable stress in speech.

The results from Experiment 1 are thus very consistent with the predictions that were made *a priori* on the basis of experiments using metrical musical perception tasks and reiterative speech tasks with participants with dyslexia. High-functioning adults with dyslexia showed difficulties in the auditory perception of rise time and difficulties in perceiving syllable stress. Individual differences in rise time perception predicted individual differences in stress perception. However, as the two spoken items to be judged were identical, the task was rather easy for all the participants. We therefore repeated the experiment using different real word tokens in the same stress perception task. Using different words increases the cognitive load of the task, as differences in segmental phonology must be ignored, making it likely that abstract stress templates must be extracted and compared. Experiment 2 therefore measures more than stress perception per se, and is conceptually more similar to the reiterative speech (DeeDee) task in requiring a more abstract stress-based comparison.

Table 5
Unique variance (R^2 change) in the syllable stress task in Experiment 1 (d') in 2-step fixed entry regression equations.

Step	Beta	R^2 change
1. Group	−.40	.16*
2. Rise time	−.56	.24**
2. Frequency	−.02	.00
2. Intensity	−.25	.06

Beta = standardized Beta coefficient; R^2 change = unique variance accounted for at each step of the 2-step fixed entry multiple regression equations.
* $p < .05$.
** $p < .01$.

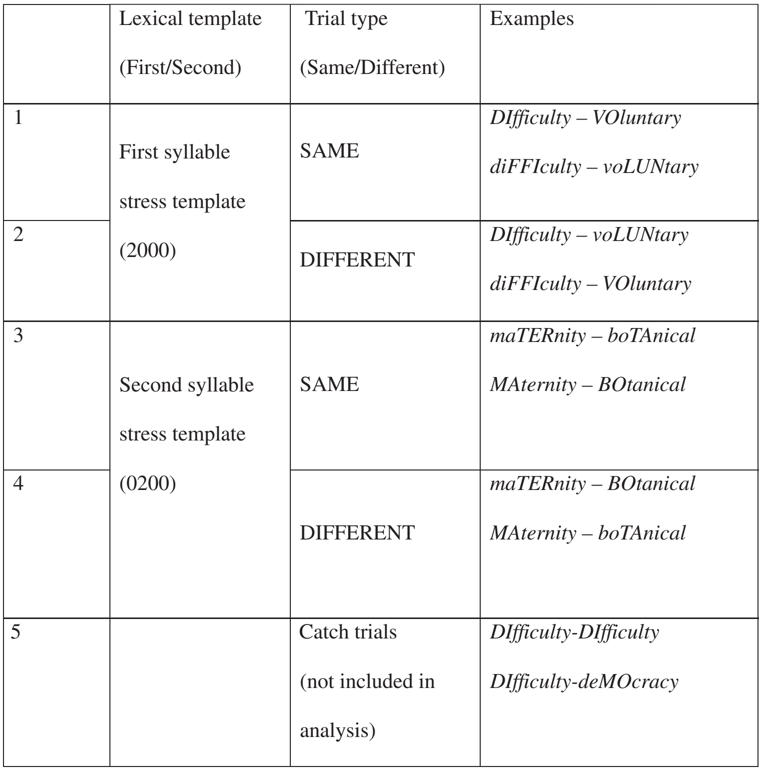
Experiment 2

Participants and tasks were as in Experiment 1, but the stress judgement task was based on pairs of two different words.

Syllable stress task

In Experiment 2, the words were 10 pairs of non-identical tokens created by pairing the 20 items from Experiment 1. Five pairs had first syllable stress templates (2000, e.g., *difficulty–voluntary*), and the other five pairs had second syllable stress templates (0200, e.g., *maternity–botanical*). This factor is referred to as First/Second. The pairs again either had the same stress (SWWW–SWWW or WSWW–WSWW) or different stress (SWWW–WSWW or WSWW–SWWW). This factor is referred to as Same/Different Judgement. Word pairs were presented in both possible orders (e.g. *difficulty–voluntary* and *voluntary–difficulty*). This resulted in a total of $10 \times 2 \times 4 = 80$ experimental trials. The experiment was again based on a $2 \times 2 \times 2$ design (Group × First/Second × Same/Different Judgement). Fig. 3 shows a schematic depiction of the design of Experiment 2, and also provides examples of the word pairs used. Word pairs were selected to have similar spoken frequencies. Appendix B provides the full list of word pairs presented.

As this second syllable stress task was substantially more difficult for participants, we added filler items containing novel pairings to discourage the use of memory strategies. These filler items comprised 20 additional easy 'catch' trials containing pairs of the same word (e.g. DIfficulty–DIfficulty as in Experiment 1), and 20 additional trials containing novel pairings of words with different lexical stress templates (e.g. DIfficulty–deMOcracy). These



	Lexical template (First/Second)	Trial type (Same/Different)	Examples
1	First syllable stress template (2000)	SAME	*DIfficulty – VOluntary* *diFFIculty – voLUNtary*
2		DIFFERENT	*DIfficulty – voLUNtary* *diFFIculty – VOluntary*
3	Second syllable stress template (0200)	SAME	*maTERnity – boTAnical* *MAternity – BOtanical*
4		DIFFERENT	*maTERnity – BOtanical* *MAternity – boTAnical*
5		Catch trials (not included in analysis)	*DIfficulty-DIfficulty* *DIfficulty-deMOcracy*

Fig. 3. Schematic depiction of design of Experiment 2.

novel pairs were included to reduce the likelihood that participants would use a strategy of relying on memory for the exact word pairs that had already been presented, rather than making judgments based on the actual stress pattern of the words. These 40 extra trials were not included in the analyses. There were thus 120 trials in total in Experiment 2, fully randomised and presented in 5 blocks of 24 trials each.

Results

Mean performance (% correct and reaction time) in each condition, and overall d' and criterion values are shown in Table 6. As can be seen, control performance on average was above 80% correct for all conditions, but the participants with dyslexia performed at a much lower level. Reaction time was again very similar across groups, and no differences by Group in response times were found in preliminary analyses. Response time is not analysed further. Paired t-tests for d' and c values revealed significant group differences for sensitivity, but not for criterion bias. Participants with dyslexia again showed a significantly lower sensitivity (d') than controls ($t(38) = 5.9$, $p < .001$). However, there was no significant difference in the response bias of both groups, indicating that neither group was more biased toward giving a 'same' or 'different' response. The d' measure from Experiment 1 was highly correlated with the d' measure from Experiment 2 ($r = 0.56$, $p < .001$).

In order to explore the effects of the experimental manipulations, a 2 × 2 ANOVA (Group × First/Second

Table 6
Group performance on the stress perception task in Experiment 2: Mean% correct, Mean RT, d' and c (*sd* in parentheses).

	% Correct		RT in ms	
	Dyslexic	Control	Dyslexic	Control
First syllable stress template (2000)				
Same judgement	64 (13.9)	88 (10.8)	2100 (674)	1783 (669)
Different judgement	59.8 (17.7)	85.3 (18.8)	2303 (765)	1832 (686)
Second syllable stress template (0200)				
Same judgement	68.3 (13.3)	86.5 (12.4)	2089 (817)	1787 (665)
Different judgement	51.8 (19.0)	82.3 (19.2)	2311 (794)	1936 (772)
d' (sensitivity)	1.2 (0.9)	3.2 (1.2)		
Criterion (bias)	0.1 (0.3)	0.0 (0.3)		

stress) was again carried out, taking d' as the dependent variable. The ANOVA showed a significant main effect of Group, $F(1, 38) = 38.1$, $p = .000$, but no significant main effect of First/Second stress, $F(1, 38) = 2.0$, $p = .161$, and no significant interaction between First/Second stress × Group, $F(1, 38) = .02$, $p = .898$. Overall, as in Experiment 1, Experiment 2 found significantly less accurate performance by individuals with dyslexia, irrespective of the stress judgement (SWWW, SWSS) required.

To explore whether individual differences in basic auditory processing contributed to individual differences in making judgements about syllable stress when two different words had to be compared, multiple regression analyses were again used. Three 2-step fixed order equations were again computed, again entering Group at Step 1 and rise time threshold, frequency threshold or intensity threshold at Step 2. The dependent variable was d'. The results are shown in Table 7. As can be seen, rise time discrimination contributes 5% of unique variance to the accuracy of judgements about syllable stress, a finding which approached significance ($p = .07$). Neither frequency discrimination nor intensity discrimination contributed unique variance (0% and 1% respectively). As d' was significantly related in the two experiments, we also present analyses for average d' in Table 7. Average d' is a measure of stress sensitivity across the two experiments combined. As Table 7 shows, rise time was the only significant predictor of individual differences in making judgements about syllable stress, even when Group was controlled as a factor.

Finally, we were interested in the relationships between performance in the stress perception tasks (assessed via d' in Experiments 1 and 2, and the average d' measure) and performance in the literacy, phonology and language measures. The full correlation matrix is shown in Table 8. Table 8 shows that prosodic sensitivity as measured by the stress perception tasks is significantly related to individual differences in reading, spelling, phonological skills and RAN. The correlations suggest that stress processing is related to both phonological and literacy performance in this sample, although the direction of causation cannot be assessed. However, it is possible to use logistic regression to predict each individual's group membership (control or dyslexic) on the basis of their performance on these different measures. Therefore, a backwards stepwise logistic regression analysis was conducted. The regression model was initialised with four predictor variables – reading, phonology (Spoonerisms), average d' across both syllable stress experiments, and rise time threshold. As will be recalled, the groups differed significantly on all four of these variables. In the backwards method, predictors that do not contribute a significant change to the likelihood ratio statistic are removed sequentially until only significant predictors remain in the model. Table 9 shows the results from this first set of logistic regressions. Only two predictors for group membership were retained in the final model – syllable stress and reading. Of these, syllable stress was the stronger predictor, contributing a larger change to the likelihood ratio statistic. In contrast, phonology and auditory perception were not retained in the model as significant predictors of group membership. Having identified syllable stress perception and reading as the strongest predictors for group membership, a second stepwise logistic regression was conducted using only these variables. Syllable stress (average d') was entered as the first step since this was the strongest predictor in the backward model. Reading was entered as the second step. As shown in Table 10, syllable stress alone correctly predicted group membership for 80% of participants. Adding reading to the regression model improved the accuracy of predictions to 87.5%. Overall, these data suggest that stress perception is a more persistent discriminator of dyslexic difficulties than phonological or auditory measures, at least when participants are high-performing and well-compensated dyslexics, as was the case for the current sample.

Table 7

Unique variance (R^2 change) in the syllable stress task in Experiment 2 (d', see 7A) and in both experiments combined (average d', see 7B) explained by the basic auditory processing measures in 2-step fixed entry regression equations.

	Beta	R^2 change
7A		
1. Group	−.69	.48***
2. Rise time	−.25	.05^{a}
2. Frequency	−.06	.00
2. Intensity	.09	.01
7B		
1.Group	−.68	.46***
2. Rise time	−.37	.10**
2. Frequency	−.05	.00
2. Intensity	.02	.00

Beta = standardized Beta coefficient; R^2 change = unique variance accounted for at each step of the 2-step fixed entry multiple regression equations.

** $p < .01$.

*** $p < .001$.

^{a} $p = .07$.

Table 8

Raw correlation matrix for Experiment 1 d', Experiment 2 d' and the average d' measure.

	Experiment 1 d'	Experiment 2 d'	Av. d'
Age	.09	.15	.14
NVIQ	.18	.33*	.31*
VIQ	.35*	.14	.22
Rise thresh	−.62**	−.53**	−.61***
Frequency thresh	−.12	−.30	−.28
Intensity thresh	−.30	−.02	−.09
Spoonerisms	.11	.57***	.50**
RAN	−.35*	−.46**	−.47**
Reading	.39*	.53**	.54***
Spelling	.27	.53**	.51**
Digit span	.32*	.50**	.50**

Note: Expt = Experiment; NVIQ = non-verbal IQ (standard score on WASI Blocks subtest); VIQ = standard score on WASI Vocabulary subtest; Reading/spelling = reading/spelling standard score on Wide Range Achievement Test, Spoonerism = No. correct on spoonerisms task, RAN = naming speed averaged across dense and sparse object RAN, Digit span = standard score on WASI digit span test.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 9
Backwards stepwise (likelihood ratio) logistic regression for participant group membership using reading, phonology, syllable stress and rise time threshold as predictors.

Step	Predictors	*B*	Exp *b*	Change in −2 log likelihood if variable removed	Model R^2 (Nagelkerke)	Overall % correct predictions (%)
1.	Stress (Av. *d*′)	−2.12*	.12	6.56*	.68	81.8
	Reading	−.13	.88	2.28		
	Rise time	.05	1.1	.76		
	Spoonerisms	.10	1.1	.20		
2.[a]	Stress (Av. *d*′)	−1.93*	.15	6.80**	.68	81.8
	Reading	−.11	.90	2.28		
	Rise time	.06	1.1	1.08		
3.[b],[c]	Stress (Av. *d*′)	−2.10*	.12	9.83**	.66	84.8[d]
	Reading	−.12*	.89	3.44		

B = regression coefficient, significance calculated using Wald statistic; exp *b* = change in odds ratio; Model R^2 = total variance accounted for by the model at each step.

* $p < .05$.

** $p < .01$.

[a] Variable removed on step 2 = spoonerisms.

[b] Variable removed on step 3 = rise time.

[c] Model $\chi^2(2) = 22.31$, $p < .001$.

[d] Correct predictions for controls = 81.3%, dyslexics = 88.2%.

Table 10
Stepwise logistic regression for participant group membership using syllable stress and reading as predictors.

Step	Predictors	*B*	Exp *b*	Model R^2 (Nagelkerke)	Overall % correct predictions (%)
1.	Stress (Av. *d*′)	−2.47**	.09	.58	80.0
2[a]	Stress (Av. *d*′)	−2.08*	.13	.69	87.5[b]
	Reading	−.16*	.86		

Note: *B* = regression coefficient, significance calculated using Wald statistic; exp *b* = change in odds ratio; Model R^2 = total variance accounted for by the model at each step.

[a] Model $\chi^2(2) = 29.22$, $p < .001$.

[b] Correct predictions for controls = 85.0%, dyslexics = 90.0%.

* $p < .05$.

** $p < .01$.

Discussion

We proposed here that very basic auditory processes may be required to perceive periodic structure in speech, following the multi-tier framework for understanding spoken language proposed by Greenberg (2006). On the basis of our prior data with children with dyslexia, we also proposed that individual differences in basic auditory processing of rise time may affect the development of metrical language processing skills such as the perception of spoken syllable stress. Given the importance of accurate prosodic perception for phonological development (Pierrehumbert, 2003), and the well-documented phonological deficits found in developmental dyslexia, we expected difficulties in stress perception in adult individuals with dyslexia. Consistent with this hypothesis, the same-different judgement task designed here to measure stress perception in adults was indeed found to be performed less accurately by adults with developmental dyslexia. This difficulty was consistent across two experiments, whether adults were making a judgement about an identical lexical item repeated twice (*maternity–maternity*), or about two different lexical items (*maternity–ridiculous*). This suggests that individuals with dyslexia are impaired in the detection of acoustic prominence in speech.

In addition, correlational analyses demonstrated that individual differences in the accuracy of stress perception were associated with individual differences in rise time discrimination, for both the "easy" (Experiment 1) and the "difficult" (Experiment 2) versions of the stress perception task, as well as for performance averaged across the two experiments (average *d*′). These relationships are consistent with data from previous studies utilising both indirect stress sensitivity paradigms (such as reiterative speech, Goswami et al., 2009), and a metrical perception paradigm involving music (Huss et al., 2010). For both reiterative speech and metrical structure in music, rise time discrimination was also found to be a significant predictor of individual differences in performance accuracy. Although participants with dyslexia in the current study showed poorer auditory discrimination of *both* rise time and pitch, only individual differences in rise time discrimination predicted stress perception. Rise time may be a more important acoustic cue to acoustic prominence than pitch (cf. Greenberg, 2006), as rise time quantifies the change in sound energy (intensity of the signal) produced by speakers as they articulate the onsets of stressed and unstressed syllables. Intensity discrimination, which was also related to accuracy in the musical metrical perception task used with children by Huss et al. (2010), was not a significant predictor of stress judgements. This makes sense, as the musical sequences in Huss et al.'s study all used the same instrument, and so only intensity and not rise time varied when notes were accented. In speech, both rise time and overall intensity will vary when syllables are accented or stressed.

Performance in the syllable stress task (average *d′* measure) was also a strong predictor of literacy, predicting group membership with 80% accuracy. This suggests that subtle speech processing difficulties in developmental dyslexia, such as the difficulty with stress perception documented here, persist into adulthood and can be stronger markers than the auditory and phonological difficulties that are markers of dyslexic difficulty in childhood. Although *a priori* there may appear to be little reason to link prosodic sensitivity and written word recognition, significant relations between stress perception and reading have been demonstrated in languages where stress is marked in the orthography, such as Greek (e.g., Protopapas & Gerakaki, 2009). Such demonstrations suggest that the perception of stress patterning in speech (the accurate detection of alternating strong and weak beats) is important for both phonological development and for acquiring literacy.

Studies are just beginning to demonstrate developmental relations between stress perception and reading acquisition, both in languages where stress is marked in the orthography (e.g., Gutiérrez-Palma & Palma-Reyes, 2007, Spanish) and in languages where it is not (Miller & Schwanenflugel, 2008, English). Even though stress is not marked by overt codes such as diacritics in English, there may be subtle orthographic cues to stress (e.g., when a syllable is written with more letters than necessary, it usually signifies that it is stressed, as in DISCUSS versus DISCUS, see Kelly, Morris, & Verrekia, 1998). Regarding phonological development, stress or prosodic patterning has been demonstrated to be an integral part of the phonological representations of individual words that are stored in the mental lexicon during infancy and early childhood (e.g., Curtin, Mintz, & Christiansen, 2005; Pierrehumbert, 2003; Vihman & Croft, 2007). During language acquisition, it appears critical that infants and children can process efficiently the temporal positions of the syllable "beats" in speech and thereby extract prosodic structure. In fact, a recent study with infants showed that statistical learning alone is a limited means of word segmentation. Johnson and Tyler (2010) studied infants' abilities to track transitional probabilities between syllables in an artificial language modelled after that used by Thiessen and Saffran (2003). The infants were aged on average 5.5 and 8 months, and two artificial languages were used, one based solely on 'words' of uniform length (CVCV), and the other based on 'words' that were either CVCV or CVCVCV. The transitional probabilities to 'word boundaries' in each language were the same. While even the 5.5-month-olds could segment 'words' in the uniform language (all CVCV), neither age group succeeded in the language with non-uniform word lengths. Johnson and Tyler (2010) noted that when artificial words are all the same length, a consistent rhythmic (periodic) cue to word segmentation is provided in addition to the transitional probability cues that are the focus of study. They suggested that more attention needed to be given to prosodic cues at the level of whole utterances in early infant word segmentation studies.

For individuals who are less sensitive to auditory cues to stress beats, in particular rise time, there may be reduced sensitivity to the rhythmic structure of speech, and this will have important consequences for developing the high-quality phonological representations of spoken words necessary for the acquisition of literacy. If a causal relationship can be established in future studies, then rhythmic and/or metrical training would be an important intervention for children with dyslexia (see Goswami, in press Huss et al., 2010, for an extended discussion). The place and role of "stress beats" (strong and weak syllables) provides temporal constraints across the different levels (syllable, word, phrase) that require functional co-ordination in speech *production* as well as speech perception (see Cummins & Port, 1998). Hence interventions addressing production as well as perception could be important. Certainly, there is ample developmental evidence that metrical structure (strong versus weak syllables) is related to how children produce words. For example, Gerken (1994) proposed a metrical template account of children's omission of weak syllables when producing multi-syllabic words. As she pointed out, during language acquisition young children are far more likely to omit weak syllables from word-initial positions than word-internal positions. The weak first syllable of a word like *giraffe* or *banana* is more often omitted than the weak second syllable of a word like *tiger*. Utilising a nonword production paradigm based on 4-syllable words, Gerken reported that while children omitted more weak syllables (45%) than strong syllables (11%) overall, their pattern of weak syllable omissions was predicted by the metrical segmentation hypothesis. For SWWS items, the first weak syllable was preserved 59% of the time, compared to 39% for the second weak syllable. However, for WSWS items, the first weak syllable was preserved 41% of the time, compared to 79% of the time for the second weak syllable. Gerken argued that young learners of English rely on metrical production templates. Infants learn rapidly from perceiving English words that they tend to begin with strong syllables, and young children apply this metrical learning to their own word productions. Our data could mean that metrical production templates would be weaker in children with developmental dyslexia.

The data presented support the view that the acoustic parameter of rise time is central to the perception of syllable stress in speech. As noted by Greenberg (2006), rise time is also important for perceiving intonational grouping because of its links with prosody. This has interesting implications for the notion that languages can be grouped into different rhythm classes, such as stress- versus syllable-timed, on the basis of different formulae quantifying consonantal and vocalic variability (e.g., Arvaniti, 2009; Grabe & Low, 2002; Ramus, Nespor, & Mehler, 1999). These formulae typically depend on durational acoustic differences, but the criteria used to place languages on a rhythmic continuum do not reflect durational variation per se, rather they depend on the extent to which a language has easily-defined prominences or accents (see Dauer, 1983, 1987; and extended discussion in Arvaniti, 2009). As rise time is the critical cue to prominence or stress accent in speech (Greenberg, 1999, 2006; Greenberg, Carvey, Hitchcock, & Chang, 2003), analyses based on rise time may help to describe stress patterning in languages that have been classically difficult to place on rhythmic continua, such as Greek, Italian and Spanish. As Arvaniti (2009) argues, rhythm does not equate to timing, as metrical structure must also be taken into consideration. She defines metrical structure as the alternation of strong and weak elements. By her account, the key acoustic factors contributing to rhythm perception in different languages are grouping and relative prominence, and durational variability plays only a small role in the creation of rhythm. Consistent with Arvaniti's linguistic argument, Huss et al. (2010) did not find that children's duration thresholds were predictive of their performance in the musical metrical task.

In their work on speech production, Cummins and Port (1998) defined rhythm in speech as the hierarchical organisation of temporally co-ordinated prosodic units. They noted that Liberman (1975) originally proposed that speech, music and dance all conformed to the "metrical organisation hypothesis", that all temporally-ordered human behaviour is metrically organised. The centrality of prosodic perception (alternating strong and weak beats) to temporally-ordered language behaviours is supported here by the strong associations found between stress perception, phonology and literacy. If human utterances are structured so that stress beats lie at privileged phases of a higher-level prosodic unit, for example marking word onsets or phrase-level information (Cummins & Port, 1998; Greenberg, 2006), then periodicity is a key organisational principle underlying phonological and intonational structure in human speech. Accordingly, an insensitivity to the auditory parameters (such as rise time) that are critical for the perception of metrical structure would be expected to affect the development of both language and literacy in children, across languages from putatively different rhythm classes (Goswami, Wang, et al., 2010). The current study provides some evidence consistent with this hypothesis.

Acknowledgments

We would like to thank our participants. This research was supported by funding from the Medical Research Council, Grant G0400574, from an EU Framework VI STREP grant, Humans as Analogy Makers, from a Leverhulme Major Research Fellowship to Usha Goswami, and by the Academy of Finland. Requests for reprints should be addressed to Usha Goswami, Centre for Neuroscience in Education, 184 Hills Rd., Cambridge CB2 8PQ, UK.

Appendix A

Word. lists

List 1	List 2
First syllable stress	*First syllable stress*
DIFFICULTY	SECONDARY
VOLUNTARY	MILITARY
COMFORTABLE	AUDITORY
ORGANIZER	CITIZENSHIP
DELICACY	LAVATORY
MONASTERY	FERTILIZER
CAULIFLOWER	DANDELION
CATERPILLAR	MERCENARY
EDUCATOR	PUNISHABLE
CATEGORIZE	PACIFIER
Second syllable stress	*Second syllable stress*
DEMOCRACY	CAPACITY
VELOCITY	RIDICULOUS
HISTORICAL	REMARKABLE
CURRICULUM	DISCOVERY
MAGNIFICENT	FACILITY
DELIVERY	NECESSITY
MATERNITY	PARTICIPANT
BOTANICAL	MANIPULATE
DEBATABLE	MIRACULOUS
HARMONICA	PISTACHIO

Appendix B

Word. pairs in Experiment 2

List 1	List 2
First syllable stress	*First syllable stress*
DIFFICULTY–VOLUNTARY	SECONDARY–MILITARY
COMFORTABLE–ORGANIZER	AUDITORY–CITIZENSHIP
DELICACY–MONASTERY	LAVATORY–FERTILIZER
CAULIFLOWER–CATERPILLAR	DANDELION–MERCENARY
EDUCATOR–CATEGORIZE	PUNISHABLE–PACIFIER
Second syllable stress	*Second syllable stress*
DEMOCRACY–VELOCITY	CAPACITY–RIDICULOUS
HISTORICAL–CURRICULUM	REMARKABLE–DISCOVERY
MAGNIFICENT–DELIVERY	FACILITY–NECESSITY
MATERNITY–BOTANICAL	PARTICIPANT–MANIPULATE
DEBATABLE–HARMONICA	MIRACULOUS–PISTACHIO

References

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica, 66*, 46–63.

Choi, J.-Y., Hasegawa-Johnson, M., & Cole, J. (2005). Finding intonational boundaries using acoustic cues related to the voice source. *Journal of the Acoustical Society of America, 118*, 2579–2587.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language & Speech, 45*, 207–228.

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and specific language impairment: A new look at an old hypothesis. *Journal of Speech, Language & Hearing Research, 50*, 1–20.

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*, 145–171.

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition, 96*, 233–262.

Cutler, A. (2005). Lexical stress. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 264–289). Oxford: Blackwell.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51–62.

Dauer, R. M. (1987). Phonetic and phonological components of language rhythm. *Paper presented at the 11th international congress of phonetic sciences* (Vol. 5, pp. 447–450). Tallinn.

Echols, C. H. (1996). A role for stress in early speech segmentation. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 151–170). Mahwah, NJ: Lawrence Erlbaum Associates.

Foxton Riviere, L.-D., & Barone, P. (2010). Cross-modal facilitation in speech prosody. *Cognition, 115*, 71–78.

Fredrickson, N., Frith, U., & Reason, R. (1997). *Phonological assessment battery* (Standardised ed.). Windsor: NFER-Nelson.

Fry, D. B. (1954). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 26*, 138.

Gerken, L. (1994). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language, 21*, 565–584.

Goswami, U. (in press). Language, music and children's brains: A rhythmic timing perspective on language and music as cognitive systems. In P. Rebuschat et al. (Eds.), *Language and music as cognitive systems*.

Goswami, U., Gerson, D., & Astruc, L. (2009). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading & Writing*. doi:10.1007/s11145-009-9186-6.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., et al. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 10911–10916.

Goswami, U., Fosker, T., Huss, M., Mead, N., & Szücs, D. (2010). Rise time and formant transition duration in the discrimination of speech sounds: The Ba–Wa distinction in developmental dyslexia. *Developmental Science* [Published online March].

Goswami, U., Wang, H. -L. S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2010). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience*. doi:10.1162/jocn.2010.21453.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.). *Laboratory phonology* (Vol. 7, pp. 515–546). Berlin: Mouton de Gruyter.

Greenberg, S. (1999). Speaking in shorthand – A syllable-centric perspective for understanding pronunciation variation. *Speech Communication, 29*, 159–176.

Greenberg, S. (2006). A multi-tier framework for understanding spoken language. In S. Greenberg & W. Ainsworth (Eds.), *Understanding speech: An auditory perspective* (pp. 411–434). Mahweh, NJ: LEA.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech – A syllable-centric perspective. *Journal of Phonetics, 31*, 465–485.

Gutiérrez-Palma, N., & Palma-Reyes, A. (2007). Stress sensitivity and reading performance in Spanish, a study with children. *Journal of Research in Reading, 30*, 157–168.

Hämäläinen, J., Leppänen, P. H. T., Torppa, M., Muller, K., & Lyytinen, H. (2005). Detection of sound rise time by adults with dyslexia. *Brain & Language, 94*, 32–42.

Hämäläinen, J., Leppänen, P. H. T., Eklund, K., Thomson, J., Richardson, U., Guttorm, T. K., et al. (2009). Common variance in amplitude envelope perception tasks and their impact on phoneme duration perception and reading and spelling in Finnish children with reading disabilities. *Applied Psycholinguistics, 30*, 511–530.

Hämäläinen, J. A., Salminen, H. K., & Leppänen, P. H. T. (in press). Basic auditory processing deficits in dyslexia: Review of the behavioural and event-related potential/field evidence. *Journal of Learning Disabilities*.

Hoequist, C. A. (1983). The perceptual centre and rhythm categories. *Language & Speech, 26*, 367–376.

Huss, M., Verney, J. P., Fosker, T., Fegan, N., & Goswami, U. (2010). Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex* doi:10.1016/j.cortex.2010.07.010.

Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science, 13*, 339–345.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159–207.

Kelly, M. H., Morris, J., & Verrekia, L. (1998). Orthographic cues to lexical stress: Effects on naming and lexical decision. *Memory & Cognition, 26*, 822–832.

Kitzen, K. R. (2001). *Prosodic sensitivity, morphological ability and reading ability in young adults with and without childhood histories of reading difficulty*. (Doctoral dissertation, University of Columbia, 2001). Dissertation Abstracts International, 62 (02), 0460A.

Klein, H. (1984). Learning to stress: A case study. *Journal of Child Language, 11*, 375–390.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency adds little. *Journal of the Acoustical Society of America, 118*, 1038–1054.

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

Liberman, M. (1975). *The intonational system of English*. Ph.D. thesis, MIT Cambridge, MA. Published by Indiana University Linguistics Club, 1978.

Lorenzi, C., Dumont, A., & Fullgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *Journal of Speech, Language and Hearing Research, 43*, 1367–1379.

Mattys, S. L., & Juscyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition, 78*, 91–121.

Mehta, G., & Cutler, A. (1988). Detection of target phonemes in spontaneous and read speech. *Language & Speech, 31*, 135–156.

Miller, J., & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly, 43*, 336–354.

Morton, J., Marcus, S. M., & Frankish, C. (1976). Perceptual centres (P-centres). *Psychological Review, 83*, 405–408.

Muneaux, M., Ziegler, J. C., Truc, C., Thomson, J., & Goswami, U. (2004). Deficits in beat perception and dyslexia: Evidence from French. *NeuroReport, 15*, 1255–1259.

Pasquini, E., Corriveau, K., & Goswami, U. (2007). Auditory processing of amplitude envelope rise time in adults diagnosed with developmental dyslexia. *Scientific Studies in Reading, 11*, 259–286.

Pierrehumbert, J. (2003). Phonetic diversity, statistical learning and acquisition of phonology. *Language & Speech, 46*, 115–154.

Port, R. (2007). How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology, 25*, 143–170.

Protopapas, A., & Gerakaki, S. (2009). Development of processing stress diacritics in reading Greek. *Scientific Studies in Reading, 13*, 453–483.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in speech signal. *Cognition, 73*, 265–292.

Richardson, U., Thomson, J., Scott, S. K., & Goswami, U. (2004). Supra-segmental auditory processing skills and phonological representation in dyslexic children. *Dyslexia, 10*, 215–233.

Rocheron, I., Lorenzi, C., Fullgrabe, C., & Dumont, A. (2002). Temporal envelope perception in dyslexic children. *NeuroReport, 13*, 1683–1687.

Slowiaczek, L. M. (1990). Effects of lexical stress in auditory word recognition. *Language and Speech, 33*, 47–68.

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

Suranyi, S., Csepe, V., Richardson, U., Thomson, J., Honbolygo, F., & Goswami, U. (2009). Sensitivity to rhythmic parameters in dyslexic children: A comparison of Hungarian and English. *Reading & Writing, 22*, 41–56.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716.

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *Journal of Physiology – Paris, 102*, 120–129.

Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading, 29*, 334–348.

Vihman, M., & Croft, W. (2007). Phonological development: Towards a "radical" templatic phonology. *Linguistics, 45*, 683–725.

Wechsler, D. (1998). *The wechsler adult intelligence scale* (3rd ed.). London: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: The Psychological Corporation.

Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading, 29*, 288–303.

Wilkinson, GS. (1993). *Wide range achievement test 3*. Wilmington, DE: Wide Range.

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

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholiguistic grain size theory. *Psychological Bulletin, 131*, 3–29.