Auditory stream segregation in dyslexic adults

P. Helenius, K. Uutela and R. Hari

Summary
Developmental dyslexia is often associated with problems in phonological processing based on, or accompanied by, deficits in the perception of rapid auditory changes. Thirteen dyslexic adults and 18 control subjects were tested on sequences of alternating tones of high (1000 Hz) and low (400 Hz) pitch, which at short stimulus onset asynchronies (SOAs) led to perceptual separation of the sound sequence into high- and low-pitched streams. The control subjects perceived the tone sequence as connected down to SOAs of 130 ms, with segregation of the streams at shorter SOAs; in dyslexic subjects the segregation occurred already at 210 ms. Auditory stream segregation has previously been shown to impair the detection of phoneme order in segments of speech sounds. The observed aberrant segregation of sound streams in dyslexic subjects might thus contribute to their difficulties in achieving awareness of phonemes or phoneme order and in the acquisition of literacy.

Keywords: dyslexia; specific language impairment; auditory processing; stream segregation; psychoacoustics

Abbreviations: SLI = specific language impairment; SOA = stimulus onset asynchrony

Introduction
Developmental dyslexia, marked by difficulties in acquiring literacy, has been associated with problems in phonological processing (Bradley and Bryant, 1983; Wagner and Torgersen, 1987; Shaywitz, 1996). Phonological processing refers to use of the sound information of spoken language and consists of phonological awareness, phonological recoding in lexical access and phonetic recoding in working memory, which all play a significant role in learning to read (Wagner and Torgesen, 1987). In line with behavioural data, functional imaging studies have revealed different patterns of cortical activation in reading-impaired individuals compared with normally reading subjects during the performance of a rhyming task (Rumsey et al., 1992; Shaywitz et al., 1998) and during the reading of isolated words (Salmelin et al., 1996; Rumsey et al., 1997).

Defects in phonological processing have been speculated to derive from a more general difficulty in the perception of rapidly changing auditory stimuli (Tallal, 1980; Reed, 1989; Tallal et al., 1993). In children with specific language impairment (SLI), the auditory defect is especially apparent, thus hindering the normal development of spoken language (Tallal and Piercy, 1973). In dyslexic children the auditory perceptual deficits are milder (Tallal, 1980), but possibly severe enough to interfere with these children’s ability to detect rapid frequency transitions in spoken consonants. Subtle deficits in language reception could interfere with the establishment of connections between the phonemes of the spoken language and the corresponding graphemes of the written language, and thus lead to difficulties in learning to read.

Previous studies have indicated that dyslexic children have no difficulties in detecting auditory stimuli which are presented in isolation or separated by a gap exceeding 400 ms (Tallal, 1980). However, when two 75-ms tones (100 and 305 Hz) were presented in rapid succession, the dyslexic children made a large number of errors both in the same–different identification task (Tallal, 1980) and in judging the temporal order of the sounds (Tallal, 1980; Reed, 1989). As the performance was impaired both in the same–different task and in the temporal order judgement task at rapid presentation rates, it was concluded that it is not the perception of temporal order but rather the sound discrimination underlying successful performance in both tasks that is deficient in dyslexic children (Tallal, 1980). Impaired discrimination of sounds has also been demonstrated in adult dyslexics using pure tones (McAnally and Stein, 1996) and frequency-modulated tones, the identification of which necessarily requires sensitivity to temporal cues (Witton et al., 1998).

As rapidly presented sounds clearly affect the perception of previous and following sounds (Bregman, 1990, 1993), the impairment of auditory processing in dyslexic children
should be even more clear with sound sequences instead of tone pairs. Accordingly, dyslexic children make more errors than control children in matching sequences composed of three to five elements; however, the impairment seems to be present even at relatively slow presentation rates (Bryden, 1972; McGivern et al., 1991). Since working memory, which may be associated with reading difficulties (Jorm, 1983; Wagner and Torgesen, 1987), certainly contributes to the results obtained, the involvement of auditory perceptual deficits in processing tone sequences has remained unsolved.

We recently studied dyslexic adults using a sequence of binaural clicks. The sequence consisted of four left-ear leading clicks followed by four right-ear leading clicks. When the clicks were presented with intervals exceeding 150 ms, the control subjects reported hearing four clicks from the left side followed by four clicks from the right side. Higher presentation rates produced an illusory perception of saltatory sound movement from left to right (Hari, 1995). In dyslexic adults the illusion persisted at significantly longer interclick intervals than in control subjects (Hari and Kiesila, 1996). Since the subjects indicated the spatial location of individual clicks only after hearing the whole sequence, this task also involves, to some extent, working memory.

The perception of successive sound sequences can, however, also be studied without loading working memory. For example, one facet of the processing speed of the central nervous system is reflected in the perceived changes of tone streams that accompany changes in stimulus presentation rate. Auditory stream segregation is a phenomenon that can occur with sound sequences consisting of alternating high- and low-pitch tones. The percept depends on both the frequency and time separation between the successive tones (van Noorden, 1975): when the frequency separation between tones is small or when the presentation rate is slow, the subjects report a connected series of tones (Fig. 1A). However, increasing the frequency separation or presentation rate results in a percept of two separate sound streams, one with higher and the other with lower pitch (Fig. 1B). The threshold of hearing sound streams as separate is called the fission boundary, and it is only slightly influenced by the rate of presentation. However, if the tones are separated by an interval of less than about 4 semitones they cannot be segregated into higher or lower streams. The threshold of single-stream percept, the temporal coherence boundary, is strongly influenced by both the frequency and the presentation rate of the stimulus. Between these two boundaries, either percept can occur according to the attentional set of the subject (van Noorden, 1975).

Auditory stream segregation can be interpreted to reflect the auditory system’s tendency to assume that a sound sequence coming from the same source does not change its properties abruptly (Bregman, 1990, 1993). Thus, with rapid presentation rates, alternating tones with adequate frequency separation are interpreted as separate sound streams arising from distinct sound sources. Segregation has an adverse effect on the perception of the temporal order of sounds (Warren et al., 1969; Bregman and Campbell, 1971). Accordingly, auditory stream segregation occurring for segments of speech sounds prevents the accurate perception of the temporal order of phoneme segments (Lackner and Goldstein, 1974; Dorman et al., 1975). However, streaming is not normally perceived during analysis of natural speech. Speech perception could thus be speculated to be based on larger units than phonemes to which stream segregation may not apply (Lackner and Goldstein, 1974). Even more importantly, formant transitions effectively bind together phonetic segments so that the temporal order of speech is normally preserved even at rapid presentation rates (Dorman et al., 1975). Formant transitions are present in, for example, stop consonants, the detection of which is impaired in dyslexic children (Reed, 1989).

As auditory perceptual deficits seem to be detectable also in adulthood (Hari and Kiesila, 1996), we tested dyslexic adults in an auditory stream segregation task to further elucidate the connection between processing of sound sequences and dyslexia. Since recent studies have indicated that speech perception deficits of language learning-disabled children can be improved by intensive training (Merzenich et al., 1996; Tallal et al., 1996), it is of great importance to gain further knowledge of the nature of the perceptual deficit underlying developmental dyslexia. Further, the well known heterogeneity of the dyslexic population (Ellis, 1985; Seymour, 1987) led us to examine the relationship between stream segregation and reading-related behavioural tasks.

Methods

Subjects

We tested 18 healthy adults (eight females; aged 20–39 years) and 13 dyslexic subjects with no history of other neurological
Auditory segregation and dyslexia

Table 1 Background of 18 control and 13 dyslexic subjects and their behavioural profiles in reading-related tests

<table>
<thead>
<tr>
<th></th>
<th>Control subjects (mean ± SD)</th>
<th>Dyslexic subjects (mean ± SD)</th>
<th>Significance (univariate F test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>26.3 ± 5.4</td>
<td>30.1 ± 7.7</td>
<td>n.s.</td>
</tr>
<tr>
<td>Educational level (years)*</td>
<td>14.4 ± 3.0</td>
<td>12.9 ± 2.3</td>
<td>n.s.</td>
</tr>
<tr>
<td>Naming (s)†</td>
<td>167 ± 27</td>
<td>98.0 ± 20</td>
<td>$P &lt; 0.0001$</td>
</tr>
<tr>
<td>Word recognition (ms)§</td>
<td>23.5 ± 5.2</td>
<td>35.7 ± 7.5</td>
<td>$P &lt; 0.0001$</td>
</tr>
<tr>
<td>Digit span (forwards)¶</td>
<td>5.8 ± 1.3</td>
<td>5.6 ± 1.4</td>
<td>$P &lt; 0.02$</td>
</tr>
<tr>
<td>Digit span (backwards)¶</td>
<td>5.9 ± 1.6</td>
<td>4.0 ± 1.2</td>
<td>$P &lt; 0.0007$</td>
</tr>
</tbody>
</table>

n.s. = not significant. *Highest level completed; † speed of reading a narrative aloud; ¶ total time to name colours, digits and letters presented in a matrix; § average time to identify Finnish words correctly from orthographically legal pseudowords in a computerized task; ‡ number of digits correctly recalled.

deficits (six females; aged 19–44 years). Two of the control and dyslexic subjects were left-handed. Informed consent was obtained from all participants. Dyslexic subjects were selected on the basis of their early history of reading and writing problems based on their own report. All but the oldest of these dyslexics had either been tested for dyslexia (10 subjects) or received special tutoring for reading difficulties during their school years (nine subjects). Some evidence of familial dyslexia (affected parent, sibling, child) was present in eight cases. The dyslexic subjects were similar to the control group for age and education level (Table 1).

**Stimuli and procedure in the behavioural tasks**

Before the psychoacoustical task all participants were tested on several reading-related behavioural tasks. However, the behavioural profiles were not used to exclude any of the subjects. Working memory span, which has occasionally been reported to be impaired in dyslexic subjects (Jorm, 1983), was tested using the standard WAIS (Wechsler Adult Intelligence Scale) procedure (Wechsler, 1955). Naming speed was used as an estimate of fluency of accessing phonological codes and was measured as the total time to name colours, digits and letters presented in a $5 \times 10$ matrix (Wolf, 1986). Oral reading speed was measured using a narrative printed on a sheet of paper. Word attack skills were tested with a computerized lexical decision task where the subjects had to indicate, as quickly as possible, if the word presented on the screen was a real Finnish word or an orthographically legal pseudoword. The words remained visible until the answer was given by pressing the response key. Word recognition speed was the mean latency of correctly recognizing a Finnish word.

As a group, the dyslexic subjects were significantly inferior to control subjects in the behavioural tasks [a repeated measures multivariate analysis of variance with the subject group as the between-subjects factor and the behavioural task as the within-subjects factor; $F(1,29) = 22.5, P < 0.0001$]. Separate between-subjects contrasts revealed that dyslexics were especially slow compared with controls in oral reading, naming and word recognition ($P < 0.0001$; Table 1). The oral reading speed of each dyslexic individual was at least one standard deviation slower than the mean of the age-matched control group. Thus, our dyslexic subjects, although with plenty of reading experience evident by their normal educational level, still lagged behind in their reading speed relative to adults without a history of reading difficulties.

**Stimuli in the stream segregation task**

Tone sequences were composed of alternating 1000- and 400-Hz pure tones, 49 ms in total duration, including 24-ms cosine rise and fall envelopes. Each sequence lasted 6 s. The sequences were presented using a Hewlett Packard workstation with a Matlab program (version 5; MathWorks, 1996), which played the sounds through headphones to both ears simultaneously at ~80 dB sound pressure level.

**Procedure in the stream segregation task**

All subjects were tested individually. Before the beginning of the test the subjects read, at their own pace, written detailed instructions of the procedure. The instructions specified the following main points: the tone sequences are composed of alternating high- and low-pitch tones; the presentation rate of the tones will be changed during the experiment; each sequence leads to either a connected or a segregated percept (illustrated on the instruction page); the subject has to press, using a mouse button, the label ‘connected’ or ‘segregated’ on the screen after hearing the corresponding sound sequence. Subjects were also instructed to follow the sound sequences as if they were connected, but as soon as the sounds irresistibly split into two streams the subjects were to report them as segregated; the aim was to find the coherence boundary (van Noorden, 1975). After the subjects had read the instructions, the sound sequences were demonstrated first at the extreme ends of the stimulus onset asynchrony (SOA; 50–800 ms) in order to clarify what the connected and segregated sequences sounded like. The subjects were also allowed to find, at their own pace, their individual temporal coherence boundaries by adjusting the SOA with a mouse-controlled ruler. Before starting the main test, the
experimenter ensured that the subject had fully understood the procedure. In the main test the subjects indicated after each sound sequence whether they had perceived the tones as connected or segregated. We used a simple adaptive one-up, one-down method, well suited and efficient for estimating the 50% level of the two forced-choice responses (Levitt, 1971). On the basis of the subject’s response the computer program automatically either shortened the SOA (after ‘connected’ answer) or lengthened it. The next sequence did not start until at least 1 s had elapsed after the subject’s response. Since there was no time pressure for the subject’s decision, the pause between separate sound sequences was actually several seconds. Each experiment comprised 30 sound sequences. The SOA of the first sequence was 600 ms, i.e. several hundreds of milliseconds longer than the mean coherence boundary of normal subjects (van Noorden, 1975). This SOA was also long enough for all dyslexic subjects to easily perceive a ‘connected’ stream. In the first 10 sequences the SOA was changed by 40 ms; during these trials the subjects were expected to approach the threshold of the coherence boundary. In the 10 next sequences the step was 20 ms; during these trials the responses were expected to stabilize near the real coherence boundary. During the last 10 trials the step was 10 ms.

The temporal coherence boundaries, defined as the mean SOA during the last 10 trials (trials 21–30), were compared between the two groups using the two-tailed t test.

**Results**

Figure 2 depicts the mean results for both subject groups in the stream segregation test. As is evident from Fig. 2A, dyslexics started to deviate from the control subjects at SOAs below 320 ms. Accordingly, the mean SOA of the first change in response type was significantly longer for dyslexics (187 ± 29 ms; mean ± SEM) than for control subjects (106 ± 7 ms; P < 0.003). In 10 of the 13 dyslexic and in all control subjects the first segregated response occurred between trials 10 and 20; in three dyslexics the first segregated response occurred within the first 10 trials. The mean coherence boundary, shown in Fig. 2B, was significantly higher for dyslexic than for control subjects (208 ± 24 versus 127 ± 9 ms; P < 0.002).

To further elucidate the connection between stream segregation and reading-related tasks, we calculated the correlation between individual subjects’ performance on stream segregation and working memory tasks, naming speed, oral reading speed and lexical decision times for words and pseudowords. We also calculated the correlation between stream segregation and lexicality effect (pseudoword recognition time/word recognition time). Both the control and dyslexic subjects recognized real Finnish words significantly faster than pseudowords (control subjects, 573 ± 21 and 652 ± 20 ms for the two types of words, respectively; two-tailed t test, P < 0.01; dyslexic subjects, 931 ± 74 and 1215 ± 115 ms; P < 0.05). The lexicality effect was significantly stronger for dyslexic than for control subjects (P < 0.02).

Control subjects behaved in a highly uniform manner in all tasks and their coherence boundaries did not correlate with any of the behavioural measures; Fig. 3 (left) shows the near-zero correlation between the controls’ coherence boundaries and their naming speed scores (Pearson r = 0.01, n.s.). The dyslexic subjects behaved differently, as illustrated in Fig. 3 (right panel): individuals who were slower in naming had a higher coherence boundary (Pearson r = 0.72, P < 0.04 using Bonferroni correction). No significant correlation was observed between the coherence boundary and other behavioural measures in dyslexic subjects, although the correlation with the strength of the lexicality effect (r < 0.70) just failed to reach significance (P < 0.08). If we accept that dyslexia represents one end of a population continuum in reading and reading-related behavioural measures (e.g. Shaywitz et al., 1992), the lack of correlation between behavioural test scores and stream segregation in the control group might derive from two factors: (i) there is not a one-to-one correspondence between phonological test scores and auditory temporal sensitivity in the range of about one standard deviation around the mean, and (ii) the relatively
small control group did not contain enough individuals near the lower end of the continuum.

Discussion
We compared the performance of dyslexic and normally reading adults in an auditory stream segregation task using alternating high- and low-pitched tones. The results unequivocally show that dyslexic adults perceived the sound sequences as segregating into two streams at significantly slower presentation rates than the control subjects did. The abnormal stream segregation in dyslexic adults may reflect prolongation of the time window during which sounds can affect the perception of previous or subsequent sounds. This interpretation is in line with the findings of our earlier study (Hari and Kiesila, 1996), which showed that the time window during which later sounds can affect the perceived location of earlier sound is extended in dyslexic adults.

Stream segregation in relation to speech processing
Previous studies have indicated that dyslexic children are impaired in identifying brief tones at rapid presentation rates (Tallal, 1980; Reed, 1989) and speech sounds that contain rapid transitions (Reed, 1989). Impaired perception of brief acoustic elements of speech has been suggested to be related to the abnormal masking effect found in SLI children (Wright et al., 1993). It is possible that, because of an extended time window of perceptual integration in dyslexic individuals, previous speech sounds could interfere with the identification of later-occurring sounds, for example by masking short transitions, and thereby lead to phonological problems. Stream segregation could relate to speech perception also by impairing the detection of phoneme order: the detection of the order of speech sounds is impaired because of streaming when the sounds do not contain formant transitions (Dorman et al., 1975), and in dyslexic children the perception of formant transitions is abnormal (Reed, 1989).

The temporal order judgement of dyslexic subjects has typically been measured with very short sound sequences, usually tone pairs, whereby the effects that are present in listening to natural connected speech may not be revealed. Furthermore, the impairment of the dyslexic and SLI children both in temporal order judgement and in same–different identification tasks has been considered to be due to problems in sound identification rather than in perceiving the temporal order of the tones (Tallal and Piercy, 1973; Tallal, 1980). In the study of Tallal and Piercy (1973), the normal children performed above chance level in the temporal order judgement task with two 75-ms tones at intervals as short as 8 ms. On the other hand, the SLI children performed at chance level at intertone intervals below 305 ms. The difference between SLI and normal children was also evident on the same–different identification task. Thus, temporal order judgement does not seem to be critical for differentiating between normal and impaired children. However, closer inspection of the results of Tallal and Piercy (1973) (see also Tallal et al., 1993) suggests (although statistical significances were not given) that in normal children temporal order judgement is impaired relative to same–different identification at intertone intervals between 30 and 150 ms. A similar difference seems to appear in SLI children at slower presentation rates, around 150–350 ms, corresponding closely to the time window during which sound streams became segregated in our dyslexic adults. Further experiments are warranted to decide whether the phonological problems of dyslexic individuals might relate, in addition to impaired identification, to difficulties in detecting phoneme order.

Temporal characteristics of stream segregation
In the present study we varied only the sound presentation rate. However, sequence duration (and the frequency difference between successive tones) may also affect stream segregation in dyslexic subjects. Stream segregation often builds up steadily over time (Anstis and Saida, 1985) and also takes some time to dissipate (Bregman, 1978). A recent study by Beauvois and Meddis (1997) further demonstrated that the effect of an induction sequence on stream segregation decays differently among listeners: in non-musicians, the induction sequence had a clear effect on stream segregation (eight 1000- and 1420-Hz tones played at SOAs of 90 ms) if the silent interval between the induction and test sequences was <3 s. In musicians, the induction sequence continued to influence the perception of the test sequence even after a silent interval of 8 s. It might well be that in dyslexic subjects also the critical time constants required for the development of segregation, as well as of its decay, could differ from those in control subjects. It would be of interest to test this hypothesis, as the difference between groups demonstrated by Beauvois and Meddis (1997) was of the order of seconds whereas the perceptual impairments thus far reported on dyslexic subjects have occurred on a time scale of tens to hundreds of milliseconds; difficulties at longer time scales would indicate a deficit in an another level of analysis of the auditory scene.

Relationship between stream segregation, naming speed and recognition of pseudowords
The coherence boundary correlated significantly with naming speed in our dyslexic adults. Naming speed reflects the fluency with which phonological information can be retrieved or accessed in long-term storage. Decreased naming speed is a specific sign of dyslexia (Wolf and Obregon, 1992), and it seems to persist into adulthood (Felton et al., 1990; Korhonen, 1995). Interestingly, naming speed has been suggested to determine the rate at which the order of tones or phonemes can still be detected (Warren, 1974), thus...
indicating a connection between dyslexia, naming speed and stream segregation.

Further support for the connection between stream segregation and phonological processing deficits was provided by pseudoword recognition times. Pseudowords were recognized more slowly than real words by both groups, but this lexicality effect was stronger for dyslexic than control subjects. Interestingly, in the dyslexic group a strong lexicality effect was modestly associated with abnormal performance on a stream segregation task. This is in line with a recent study by Witton et al. (1998) demonstrating that poor non-word reading ability is associated with impaired auditory temporal perception in dyslexic adults. The dual-route models of reading (e.g. Coltheart, 1978; Ellis and Young, 1988; Coltheart et al., 1993; Bookheimer et al., 1986) assume that the meaning of the word can be accessed either through laborious grapheme-to-phoneme conversion or, if the word is familiar, by direct activation of the meaning of the word through lexical access based on the visual form of the word. It has been suggested that dyslexic individuals who have marked problems in phonological processing rely more on the direct, lexical route (Campbell and Butterworth, 1985; Snowling et al., 1986). On the other hand, individuals with difficulties in accessing lexical representations would use grapheme-to-phoneme conversion on accessing the meaning of a word (Hanley et al., 1992; Castles and Coltheart, 1996). As pseudowords are not stored in the lexicon, they apparently have to be read through the phonological route (for opposing views, see e.g. Kay and Marcel, 1981). As grapheme-to-phoneme conversion is laborious in phonologically impaired dyslexics, pseudoword reading should take an excessively long time, as was evident in many of our dyslexic subjects.

To summarize, the present study indicates that those dyslexic adults who are slow in accessing phonological information also show clear abnormalities in processing rapidly presented tone sequences. Such an auditory deficit could signal a prolonged perceptual time window during which sounds can affect percepts of previous and subsequent sounds. This could lead to impaired detection of phonemes or phoneme order, thereby hampering the normal development of reading and writing.

Acknowledgements
We wish to thank Dr Riitta Salmelin, Dr Elisabet Service, Professor Veijo Vиру and three anonymous reviewers for their valuable comments on earlier versions of this manuscript. This study was, supported by the Academy of Finland and the Sigrid Juselius Foundation.

References


