ELSEVIER

Cognition 59 (1996) 275–306

COGNITION

Analysis of perceptual confusions between nine sets of consonant–vowel sounds in normal and dyslexic adults

P.L. Cornelissen[a,*], P.C. Hansen[a], L. Bradley[b], J.F. Stein[a]

[a] *Physiology Department, Oxford University, Parks Road, Oxford, OX1 3PT, UK*
[b] *Experimental Psychology Department, Oxford University, South Parks Road, Oxford, OX1 3UD, UK*

Received 14 September 1994, final version accepted 27 September 1995

Abstract

It is widely accepted that most developmental dyslexics perform poorly on tasks which assess phonological awareness. One reason for this association might be that the early or "input" phonological representations of speech sounds are distorted or noisy in some way. We have attempted to test this hypothesis directly. In Experiment 1, we measured the confusions that adult dyslexics and controls made when they listened to nine randomly presented consonant–vowel (CV) segments (/ba/, /da/, /pa/, /tʃa/, /fa/, /ʃa/, /la/, /wa/ & /ja/) under four conditions of increasing white noise masking. Subjects could replay stimuli and were under no obligation to respond quickly. Responses were selected with a computer mouse from a set of nine letter-strings, corresponding to the auditory stimuli, presented on a VDU. While the overall pattern of confusions made by dyslexics and controls was very similar for this stimulus set, dyslexics confused /tʃa/ with /ʃa/ and /pa/ with /fa/ significantly more than did controls. In Experiment 2, subjects heard each stimulus once only and were forced to respond as quickly as possible. Under these timed conditions, the pattern of confusions made by dyslexics and controls was the same as before, but dyslexics took longer to respond than controls. The slower responses of dyslexics in Experiment 2 could have arisen because: (a) they were slower at processing the auditory stimuli than controls, (b) they had worse visual pattern memory for letter strings than controls, (c) they were slower than controls at using the computer mouse. In Experiments 3, 4 and 5 subjects carried out control tasks which eliminated each of these possibilities and confirmed that the results from the auditory tasks

* Corresponding author. Present address: Psychology Department, Ridley Building, Newcastle University, Newcastle upon Tyne NE1 7RU.

SSDI 0010-0277(95)00697-4

genuinely reflected subjects' speech perception. We propose that the fine structure of dyslexics' input phonological representations should be further explored with this confusion paradigm by using other speech sounds containing VCs, CCVs and VCCs.

1. Introduction

1.1. General background

Despite adequate educational opportunity, many children fail to acquire competent reading skills. A surprising number of these poor readers have normal or above normal abilities in other areas, so their reading problems are unexpected. Rutter and Yule (1975) attempted to capture this sense of unexpected reading failure by using statistical criteria. They defined children as specifically retarded readers (i.e., developmental dyslexics) if their reading ability was significantly lower than that predicted on the basis of age and IQ. Stanovich (1991) has criticised IQ discrepancy measures of developmental dyslexia because of the fact that poor reading ability tends to be correlated with poor performance on IQ tests. Consequently, some authors have suggested that a mismatch between reading and spoken comprehension might provide a better measure (Gough & Tunmer, 1986). Nevertheless, the majority of published research has used some version of the IQ discrepancy measure to define developmental dyslexia. For the sake of consistency, therefore, we have used the same definition.

One approach to studying developmental dyslexia has been to examine dyslexics' reading behaviour in the light of dual-route models of skilled reading (Castles & Coltheart, 1993). Skilled readers can make use of two procedures, one often referred to as the *lexical route* and the other as the *sublexical route*. Reading aloud via the lexical route involves retrieving, from a mental lexicon, the phonological form appropriate to a particular orthographic stimulus. The lexical system is thought to be mainly responsible for handling irregularly spelled words (e.g., yacht). Reading aloud via the sublexical route involves using correspondence rules between orthographic and phonological segments to assemble appropriate pronunciations of words. The sublexical system is thought to be primarily responsible for dealing with unfamiliar and nonsense words (e.g., polmex). Empirical support for dual-route models of skilled reading comes from studies of individuals who have acquired dyslexia as a result of brain damage. Such patients are characterized by the fact that they may suffer selective loss of one or other subcomponent of the reading process (Patterson, Marshall, & Coltheart, 1985; Shallice & Warrington, 1980). Moreover, it has been claimed that dual-route computational models of reading provide a better fit to the reading error data from acquired dyslexics than do single route models (Besner, Twilley, McCann, & Seergobin, 1990), thus providing further support for their validity.

Castles and Coltheart (1993) argue that learning to become a skilled reader *must* involve acquiring the dual-route system in some sense. Therefore if an individual has a particular difficulty in acquiring either the lexical or the sublexical procedure, his or her reading pattern will reflect this: the individual will be unusually poor at reading either irregular or nonsense words. These authors studied the lexical (irregular word reading) and sublexical (nonsense word reading) skills of 56 developmental dyslexics and 56 controls. The control subjects were used to derive age-dependent norms for irregular and nonsense word reading. Castles and Coltheart (1993) showed that 75% of their dyslexics' irregular word reading was in the aberrant range – that is, a score which fell below the 10th percentile for controls. A similar percentage (72%) of the same group of dyslexics had abnormally low scores for nonword reading. While most dyslexics were poor at reading both irregular and nonsense words, Castles and Coltheart found individuals who had selective difficulty with reading either nonsense or irregular words. Therefore they concluded that two subtypes of developmental dyslexic could be discriminated on the basis of their irregular and nonsense word reading abilities.

Castles and Coltheart's results confirm many other groups' findings: the majority of dyslexics have difficulty with nonsense word reading, that is, the accurate application of grapheme–phoneme correspondence rules (Snowling, 1980; Olson, 1985; Kochnower, Richardson, & DiBenedetto, 1983). There is a large body of evidence to suggest that difficulty in dealing with grapheme–phoneme correspondences is associated with an underlying deficit in dyslexics' phonological awareness (Goswami & Bryant, 1990). This term refers to knowledge about the segmental nature of speech. It extends from the level of the syllable to more difficult and abstract concepts such as onset-rime boundaries and phonemes (in "bat", the onset is "b-" and the rime is "-at") (Treiman & Zukowski, 1991). Indeed, as far as normal reading development is concerned, a minimal level of phonological awareness is thought to be a necessary though not a sufficient condition for the development of efficient visual word recognition (Tunmer, 1989). Thus a number of studies have shown that phonological awareness in pre-school children predicts a successful outcome in reading development (Bradley & Bryant, 1985; Wagner & Torgeson, 1987).

1.2. The present study

Research has shown that, on average, groups of developmental dyslexics perform poorly on a variety of tasks which assess phonological awareness (Snowling, 1991). While this is a strong association, it should be remembered that not all dyslexics fit this pattern (e.g., Castles & Coltheart, 1993; Seymour, 1986). It is argued that poor phonological awareness causes many dyslexics to experience severe difficulties in mastering phonological rules relating spelling to sound (grapheme–phoneme correspondences) which leads in turn to their difficulties with reading nonwords. This relationship holds even when dyslexics are compared to younger children of the same reading age (Snowling & Rack, 1991).

Precise details of the nature of dyslexics' poor phonological awareness remain unclear. In part this is because different phonological awareness tasks demand more or less explicit awareness of phonological units (Cataldo & Ellis, 1988), as well as placing variable loads on memory. But different tasks also assess phonological units of difference size – some tasks require phoneme counting, while rhyme oddity tasks merely require comparisons of onsets and rimes. There is also evidence that the location of the phonological deficit may vary between individual dyslexics. For example, Snowling, Stackhouse, and Rack (1986) identified single cases where the dyslexic children had problems with either *input* (reception) or *output* (production) phonology. These authors refer to input phonology as the process responsible for "the registration of incoming auditory stimuli" the failure of which might lead to problems with auditory discrimination. Output phonology "refers to the stage where the articulatory instructions for pronunciation are registered. Deficits may occur either in assembling or in retrieving output phonological codes". In this paper, we focus on input phonology. Input phonological representations have been studied in the past using categorical perception tasks, though as we shall argue below, a better test would be a more open-ended paradigm.

In a typical categorical perception task, subjects might be asked to identify stimuli from a continuum of synthetic sounds which range smoothly from one end point (e.g., /ba/) to another (e.g., /da/). The stimuli are designed to present subjects with sufficient cues to distinguish the place of articulation. So, for example, the series from /ba/ to /da/ might change the starting frequency of the second and third formants from 800 and 1800 Hz respectively to 1700 and 2600 Hz in equal logarithmic steps.

At least six studies have been published in which categorical perception of speech sounds was investigated in normal and dyslexic subjects using synthetic CV continua (/ba-/da/ and /da/-/ga/ which differ in the frequency of the second formant transition; /bath/-/path/ and /sa/-/sta/ which differ on voice onset time) Read, 1989; Brandt & Rosen, 1980; Godfrey, Syrdal-Lasky, Millay, & Knox, 1981; Manis, McBride, Seidenberg, Doi, & Custodio, 1993; Steffens, Eilers, Gross-Glenn, & Jallad, 1992; Werker & Tees, 1987). With the exception of the /sa/-/sta/ continuum, these studies focused on the categorical perception of stop consonants. In all of these studies, the fact that dyslexics exhibited categorical perception meant that they could clearly discriminate between the consonants studied.

However, in five out of the six studies, the slopes of the boundary curves for the dyslexic subjects were less steep than normal.

A failure to replicate these results was reported by Brandt and Rosen (1980). However, the number of subjects in their experiment was small. Brandt and Rosen used four children in each of four groups, which was only a third the number of subjects used in the other five studies above. Low subject numbers in this study may have presented a particular problem if their dyslexics happened not to have had particularly poor phonological awareness. Manis et al. (1993) showed that the extent of dyslexics' poor performance on their /bath/-/path/ continuum depended on whether they also performed poorly on phonological awareness tasks (they used phoneme deletion, e.g. say /flin/ without /f/, and phoneme position analysis, e.g. what sound comes before /p/ in /frimp/). A related finding in normal children is that when they are given the rhyme "oddity" phonological awareness task (i.e., detecting that *bat* is the odd item in 'bat pad had mad'), they make more errors if the odd word is phonetically similar to the non-target words (Snowling, Hulme, Smith, & Thomas, 1994).

The demonstration of weaker categorical boundaries for certain speech contrasts raises the possibility that (some) dyslexics might experience greater confusion between acoustically similar phonemes than normals do. Associating speech sounds with appropriate sequences of alphabetic tokens would be particularly difficult if the discriminability between (some) sounds was reduced, thereby leading to difficulties with grapheme–phoneme correspondences. Put another way, it is conceivable that the early, or "input", phonological representations for speech sounds might be perturbed or "noisy" in developmental dyslexics. While this is a plausible hypothesis (ideas like this were originally expressed by Shankweiler & Liberman, 1972), we argue that categorical perception tests using only two or three consonant contrasts may not be a very sensitive way to look for potential distortions of this kind. After all, the kind of phonological representation that would be useful for parsing continuous speech is likely to contain a large number of features. Therefore, our alternative approach has been to measure the confusions that dyslexic and normal subjects made when they listened to nine randomly presented CV segments (/ba/, /da/, /pa/, /tʃa/, /fa/, /ʃa/, /la/, /wa/ & /ja/) under four conditions of increasing white noise masking (based on Miller & Nicely, 1955). We suggest that measuring confusions between a larger set of speech sounds, covering three classes of consonants, is a more sensitive tool for revealing potential sources of phoneme confusion in dyslexics than the categorical perception tasks described above. Our approach allowed us to look for confusions not only within, but also between consonant categories. Moreover, like Miller and Nicely, we were able to measure how efficiently dyslexics could detect information about different articulatory features (e.g., placing and voicing) in comparison with controls.

EXPERIMENT 1

2. Objective

In the first experiment we aimed to measure the accuracy with which dyslexics and normals could identify spoken CVs in the presence of increasing white noise masking and without time constraints.

3. Methods

3.1. Subjects

Ten dyslexic adults (4F, 6M) and 10 controls (4F, 6M) took part in this study. A *T*-test comparison showed that the mean chronological ages of the two groups were not significantly different. Using Annett's handedness inventory (Annett, 1985), 8 dyslexics showed a consistent pattern of handedness (7 right- and 1 left-handed) and 2 an inconsistent pattern. Six of the controls showed consistent handedness (all were right-handed) and 4 an inconsistent pattern. The adult dyslexics had all previously been diagnosed by educational psychologists on the basis of a significant discrepancy between general ability and written language skills. Subject characteristics are summarized in Table 1.

3.2. Audiometric and psychological testing

We used a Madsen Electronics OB822 audiometer to measure pure tone thresholds for air conduction in both ears at 250, 500, 1000, 2000, 4000 and 8000 Hz. No subject had a pure tone hearing loss in excess of 25dB HL. We averaged together the responses over 500–4000 Hz for left and right ears separately. This is the most important frequency sensitivity range for speech perception. As can be seen in Table 1, there was marked variability in the audiometry responses. In fact we found it very hard to find adult subjects who did not have some hearing loss, usually at frequencies above 1000 Hz. However, *T*-test comparisons for mean left and right ear losses showed no significant differences between dyslexics and controls.

All subjects were given five subtests from the WAIS-R IQ test battery: Vocabulary, Similarities, Object assembly, Block design and Digit span. As Table 1 shows, with the exception of digit span (see Jorm, 1983; Siegal & Linder, 1984), *T*-test comparisons showed that there were no significant differences in mean performance on the audiometry and IQ subtests between dyslexics and controls.

Finally, each subject was asked to read aloud all the items from the Schonell reading accuracy test (Schonell, 1950) as well as a list of 26

Table 1
Characteristics and performance of dyslexics and controls on the reading tests, audiometry and IQ subtests

	Dyslexics $n = 10$ mean (SD)	Controls $n = 10$ mean (SD)	p value
Age (years:months)	29.1 (9:5)	28:4 (4:10)	>.5
Total time taken to read 26 nonwords (s)	47.3 (20.2)	23.9 (6.3)	<.01
Total errors made reading 26 nonwords	7.3 (2.9)	3.1 (1.8)	<.01
Total time taken to read 100 real words (Schonell test) (s)	117.3 (49.1)	60.6 (6.9)	<.01
Total number of errors made on 100 real words (Schonell test)	6.8 (5.1)	2.3 (2.2)	.01
Left ear average hearing loss in dBHL for 500, 1000, 2000 and 4000 Hz tones	2.6 (1.5)	3.5 (7.1)	>.5
Right ear average hearing loss in dBHL for 500, 1000, 2000 and 4000 Hz tones	4.1 (3.3)	3.6 (2.2)	>.5
WAIS-R Digit span	10.1 (2.4)	12.9 (3.2)	<.05
WAIS-R Vocabulary	15.3 (1.8)	15.1 (1.9)	>.5
WAIS-R Similarities	13.8 (0.9)	13.5 (2.4)	>.5
WAIS-R Block design	15.4 (2.8)	14.5 (2.9)	.5
WAIS-R Object assembly	13.7 (3.6)	13.1 (2, 3)	>.5

nonwords (Snowling, personal communication). When the reading tests were being explained to the subjects, it was emphasized that even though they were being timed, they were encouraged to read words as accurately as possible. Fig. 1 shows plots of the number of correct responses against the time it took to complete the Schonell and nonword lists respectively. Fig. 1

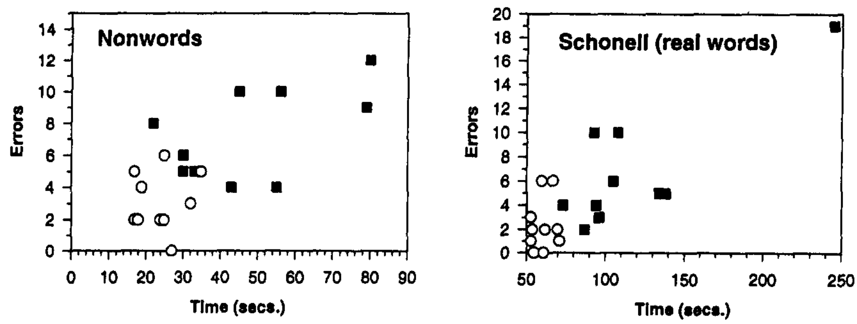


Fig. 1. Two plots of the total number of reading errors that each subject made against the time (s) it took them to complete the reading lists. The first plot is for the nonword reading test and the second plot for the real word reading test. See text for details. In each case the control and dyslexic data points are indicated by empty circles and filled squares respectively.

and Table 1 show that these dyslexic adults had persistent reading difficulties with both real and nonsense words.

3.3. Experimental stimuli

Real speech sounds were collected in a sound-proofed room using a Tucker-Davis Technologies AD/DA card with peripheral hardware connected to an IBM PC. We recorded speech with an A/D sampling rate of 22 kHz which was just over two times the frequency of our anti-aliasing filter (low-pass, 10 kHz cut-off). A male speaker sat 20 cm in front of an Audio-Technica PRO4L microphone while simultaneously listening to a pure tone (110 Hz) over headphones. With extensive practice, this allowed the speaker to utter multiple examples of each CV segment in a monotone of constant pitch and at a consistent sound pressure.

We used CSRE4.0 (Canadian Speech Research Environment version 4.0) to analyse and edit the speech recordings off-line. We selected 10 examples of each of the CV segments: /ba/, /da/, /pa/, /tʃa/, /fa/, /ʃa/, /la/, /wa/ & /ja/. We only chose samples in which the amplitude of the vowel component fell within an RMS value of 2.5 ± 0.25 V. Each of the 90 samples was then trimmed to a length of 250 ms and an additional 250 ms of silence added to its front end. Next, four levels of white noise mask were added to each of the speech samples to generate our experimental stimuli. This process generated experimental stimuli each of which lasted 5000 ms. During the first 50 ms, the white noise mask was ramped up to a constant level which was maintained for a further 200 ms. Then the speech embedded in the white noise mask began and continued for the remaining 250 ms of the stimulus. The respective signal/noise (S/N) ratios for the CV stimuli embedded in noise were typically 19, 0, −2 and −3 dB.

3.4. Procedure

Subjects sat in a quietened room wearing a pair of Sennheiser HD520II headphones. Stimuli were played back monophonically at a maximum sound pressure of 87 dB at the tympanum (approx. 70 dB at the pinna), measured with a closed ear system (Pralong & Carlisle, 1994). We used the CSRE4.0 software to control our experiment which was divided into three blocks. In each block subjects were presented all 10 examples of the nine different CV stimuli, under each of the four masking conditions. The order in which stimuli were presented was randomized. This procedure generated 360 trials per subject per block (1080 in total). At the start of each block, all the 19 dB S/N stimuli were presented first, then all of the 0 dB S/N stimuli, and so on, finishing with all of the −3 dB S/N stimuli.

Responses were recorded using a multiple-choice window presented on a VDU with mouse control. The nine possible responses were represented on a screen in a 3 × 3 array of boxes each of which contained one of nine letter strings (ba, da, pa, cha, fa, sha, la, wa & ya). Subjects sat 50 cm from the VDU, so that the letter "a", for example, subtended 1.1° × 1.1°. The box array was contained in an area which subtended about 22° horizontally ×16° vertically. The positions of the letter strings in the 3 × 3 array of boxes were randomized. These random positions were changed four times in every block, that is, every time a set of stimuli at the next S/N ratio was presented. This controlled for any tendency for subjects to prefer clicking on particular box locations. Once a subject had decided which of the nine stimuli he/she had heard, he/she clicked the mouse cursor on the appropriate response box. Visual feedback was provided. If a response was correct then the box which the subject had selected flashed green. If the response was incorrect, then the box which the subject should have selected flashed red. The computer logged both the stimulus and whether the response was correct or not.

The time course of a typical trial was as follows. The subjects first heard a brief warning beep (140 ms) whose onset and offset were linearly ramped. 100 ms later the speech stimulus was presented. At this point the subject could either make a response, or replay the stimulus as many times as they wished. During initial instruction, subjects were told that two or three replays were often helpful in deciding what they had heard, but many more than that could potentially add further uncertainty. Once the subject responded by clicking on one of the response boxes, feedback was presented for 250 ms. This was followed by a further 100 ms gap before the next warning beep.

4. Results

4.1. Confusion data

Table 2 shows the raw data from Experiment 1 as confusion matrices. In these tables the stimuli presented are indicated by the consonants listed vertically in the first column to the left of each matrix. The particular order in which the consonants are presented is based on a description of the manner of their production. Thus they have been grouped into three subsets: (a) stops, (b) affricate and fricatives, (c) liquid and semi-vowels (see Appendix A for definition of these features). The CVs which subjects gave as responses are indicated horizontally across the top of each matrix. The number in each cell is the frequency that each stimulus–response pair was observed summing across all subjects in each of the dyslexic or control groups. The number of correct responses can be obtained by totalling the frequencies along the main diagonal. Off-diagonal elements represent CV confusions. Row sums give the frequencies of presentation for each CV (i.e., 300 in each case). Column sums give the frequencies with which each CV was given as a response.

Inspection of the matrices in Table 2 shows that all subjects made more confusions with increasing white noise masking which we would expect from Miller and Nicely (1955). The manner of articulation, as defined by the three subsets above, was not a very good predictor of confusability; if it were, most of the confusions would have fallen within the boxes marked with heavy outlines (Table 2). It is also clear that very sonorant consonants, that is, the liquid and the two semi-vowels, were readily distinguished from the obstruents because there are relatively few confusions in the shaded boxes (Table 2). Most confusions were made between stops and affricative/fricatives. Overall, the most striking finding on casual inspection was how similar the dyslexics' performance was to that of the controls.

To quantify these descriptive findings we carried out a repeated measures three-way ANOVA on the numbers of confusions that subjects made (i.e., Confusions (72 levels) × Mask (4 levels) × Group (2 levels)). In this analysis we concentrated on the off-diagonal elements of the confusion matrices and discarded the diagonal elements. Because we were counting the frequency of confusions, the appropriate error distribution for these data is binomial. Therefore we transformed the confusion frequencies with the arcsin transformation and used these as the dependent variable (Snedecor & Cochran, 1967). The main effects of *confusion* and *mask* were significant, $F(71, 1278) = 65.69$, $p < .0001$ and $F(3, 54) = 461.6$, $p < .0001$, while the main effect of *group* was not significant, $F(1, 18) = 2.31$, $p < .1$. Therefore we confirmed the qualitative impression that both groups of subjects were producing essential similar patterns of confusions. However, we also found a significant Group × Confusion interaction, $F(71, 1278) = 2.79$, $p < .0001$). We carried out post hoc multiple comparisons to elucidate the nature of this

Table 2

Experiment 1: confusion matrices. In these matrices, consonants are grouped according to the manner of their articulation: (a) stops (/ba, da, pa/), (b) affricate and fricatives (/fa, ʃa, tʃa/), (c) liquid and semi-vowels (/la, wa, ja/). See Appendix A for definitions of these features. The heavily outlined boxes represent the confusions that subjects made based on manner of articulation. The shaded boxes represent the confusions that subjects made between sonorant consonants (/la, wa, ja/) and obstruent consonants (/ba, da, pa, fa, ʃa, tʃa/). See Appendix A for the definition of sonority. Signal to noise ratio (S/N) was defined as:

$$S/N\ \text{ratio} = 10.\log_{10}\frac{(\text{RMS signal})^2}{(\text{RMS noise})^2}$$

NORMALS

19DB S/N

	b	p	d	f	ch	sh	w	l	y
b	281	3	1	14	0	1	0	0	0
p	1	298	0	0	1	0	0	0	0
d	0	0	298	0	0	0	0	2	0
f	8	4	1	286	0	0	0	0	1
ch	0	0	1	0	297	2	0	0	0
sh	0	1	1	0	26	272	0	0	0
w	0	0	0	0	0	0	292	0	8
l	0	1	2	1	0	0	0	295	1
y	0	0	0	0	0	0	0	1	299
	290	307	304	301	324	275	292	298	309

0DB S/N

	b	p	d	f	ch	sh	w	l	y
b	131	44	8	112	2	3	0	0	0
p	9	249	12	12	10	5	0	2	1
d	1	0	93	0	130	75	0	0	1
f	10	120	1	151	5	11	2	0	0
ch	2	36	43	9	125	82	2	1	0
sh	0	1	16	15	75	192	1	0	0
w	0	0	0	0	0	0	294	0	6
l	0	0	5	0	2	1	2	151	139
y	0	0	0	1	1	0	1	3	294
	153	450	178	300	350	369	302	157	441

-2DB S/N

	b	p	d	f	ch	sh	w	l	y
b	156	8	16	107	4	6	0	2	1
p	4	253	8	22	7	4	1	0	1
d	5	0	92	1	85	81	1	1	34
f	5	56	5	159	26	48	0	1	0
ch	3	22	41	16	96	107	0	7	8
sh	0	2	13	21	99	164	0	1	0
w	0	0	1	0	0	0	288	2	9
l	0	0	11	1	1	1	1	194	91
y	0	0	0	3	2	3	1	29	262
	173	341	187	330	320	414	292	237	406

-3DB S/N

	b	p	d	f	ch	sh	w	l	y
b	94	51	33	74	16	18	0	5	9
p	13	152	17	38	25	20	2	25	8
d	2	1	74	3	74	93	16	6	29
f	14	46	3	163	11	54	5	3	1
ch	5	54	31	31	67	85	1	21	5
sh	3	38	11	60	47	139	1	1	0
w	1	0	2	0	1	1	270	6	19
l	1	3	17	1	1	6	21	169	81
y	0	4	6	1	14	15	12	53	195
	133	349	194	371	256	431	330	289	346

DYSLEXICS

19DB S/N

	b	p	d	f	ch	sh	w	l	y
b	284	8	0	7	0	0	1	0	0
p	1	297	0	1	1	0	0	0	0
d	0	1	299	0	0	0	0	0	0
f	10	17	2	271	0	0	0	0	0
ch	0	0	1	0	291	8	0	0	0
sh	0	0	1	1	67	230	0	1	0
w	0	0	0	0	0	0	297	0	3
l	0	1	0	0	0	0	0	299	0
y	0	0	0	0	0	0	0	0	300
	295	324	301	300	359	238	298	300	303

0DB S/N

	b	p	d	f	ch	sh	w	l	y
b	80	61	4	146	5	4	0	0	0
p	2	275	3	7	12	1	0	0	0
d	0	1	91	0	126	78	2	0	2
f	8	174	3	102	6	6	1	0	0
ch	2	23	35	12	162	63	1	1	1
sh	2	4	22	8	149	112	0	1	2
w	0	0	0	0	0	0	287	0	12
l	0	0	10	0	0	1	4	118	167
y	0	0	0	0	0	0	0	2	298
	94	538	168	268	460	265	295	122	482

-2DB S/N

	b	p	d	f	ch	sh	w	l	y
b	132	29	10	116	8	4	0	1	0
p	2	272	1	9	8	4	3	1	0
d	1	0	77	1	125	74	4	1	17
f	13	112	1	107	35	31	1	0	0
ch	1	36	33	10	132	72	1	5	10
sh	4	5	8	5	170	107	0	1	0
w	0	0	0	0	1	0	293	0	6
l	0	0	9	0	1	0	2	195	93
y	0	1	4	0	2	5	2	41	245
	153	455	143	248	482	297	306	245	371

-3DB S/N

	b	p	d	f	ch	sh	w	l	y
b	75	60	33	90	22	12	5	1	2
p	13	151	15	37	36	23	7	3	15
d	5	1	66	2	105	72	26	2	21
f	19	69	4	142	25	30	5	2	4
ch	7	35	30	34	96	74	3	19	2
sh	5	46	12	56	95	77	3	2	4
w	2	0	0	0	0	0	278	3	17
l	0	1	16	2	3	5	33	154	86
y	2	1	9	6	10	10	18	55	189
	128	364	185	369	392	303	378	247	334

interaction. Table 3 shows the five CV confusions at each masking level which best discriminated dyslexics from controls. At each S/N level ʃa/tʃa was always the confusion which best differentiated the groups. The next commonest confusion was fa/pa at all masking levels except −3 dB.

Bonferroni multiple *T*-test comparisons showed that the ʃa/tʃa confusion significantly discriminated our two groups of subjects at all masking levels ($\alpha = 0.01$), while there was only a trend in this direction for the fa/pa confusion. Thus, while the overall pattern of CV confusions was similar for both groups of subjects, dyslexics tended to make a specific excess of the ʃa/tʃa confusion.

Finally, we wanted to be sure that the variability we found in subjects' audiometry responses did not confound the speech perception results. To do this, we used the median values of the left and right ear averaged audiometry thresholds to split the 20 subjects into two groups; that is, "better" and "worse" hearers. For each subject we used the confusion matrices to calculate the sum of the diagonal elements across all masking levels. This provided a summary measure for overall performance in the

Table 3
Experiment 1: the 5 CV confusions at each masking level which best discriminate between dyslexics and controls

S/N ratio (dB)	Confusion	Dyslexics max. = 300	Controls max. = 300	Dyslexics – controls
19	sha/cha	67	26	41*
	fa/pa	17	4	13
	ba/fa	7	14	−7
	cha/sha	8	2	6
	wa/ja	5	8	−3
0	sha/cha	149	75	74*
	fa/pa	174	120	54*
	ba/fa	146	112	34
	la/ja	167	139	28
	cha/sha	63	82	−19
−2	sha/cha	170	99	71*
	fa/pa	112	56	56*
	da/cha	125	85	40
	cha/sha	72	107	−35
	ba/pa	29	8	21
−3	sha/cha	95	47	48*
	da/cha	105	74	31
	fa/sha	30	54	−24
	fa/pa	69	46	23
	da/sha	72	93	−21

* $p < .01$.

speech perception task. Finally we made T-test comparisons between "better" and "worse" hearers for the performance measure of speech perception. Left and right ear thresholds were analysed separately; neither comparison was significant, $t(18) = -0.4$, $p = .7$ and $t(18) = -0.2$, $p = .8$ respectively. Therefore, we felt justified in excluding subjects' pure tone hearing thresholds from further analysis of the speech perception results.

4.2. Articulatory features and transmitted information

We wanted to know what kinds of phonetic cues our subjects may have been using to discriminate between the CV stimuli. Miller and Nicely (1955) carried out this kind of analysis by recoding their confusion matrices for the phonetic features voicing, nasality and frication as two-part codes. They also included place which they defined as a three-part code. Miller and Nicely then quantified the degree to which information about each feature code was detected by subjects. The phonetic features they examined and the method for calculating transmitted information ($T(x; y)$) are described in Appendices A and B.

In the analysis of our data we decided not to make a priori judgements about which feature codes we should use to calculate $T(x; y)$. Instead, we calculated $T(x; y)$ for all possible three part codes for both dyslexics and controls, at each masking level. We then identified which of these codes accounted for the maximum transmitted information. We assumed that this would best reflect the features or combinations of features which subjects detected most efficiently. The results are given in Table 4, which shows how

Table 4
Experiment 1: transmitted information for three-part codes comparing dyslexics with controls

	Dyslexics		Controls	
	Code	$T(x; y)$ (bits)	Code	$T(x; y)$ (bits)
19 dB S/N	/b, f, p/ /ch, d, sh/ /l, w, j/	1.554	/b, f, p/ /ch, d, sh/ /l, w, j/	1.540
0 dB S/N	/b, f, p/ /ch, d, sh/ /l, w, j/	1.316	/b, f, p/ /ch, d, sh/ /l, w, j/	1.284
−2 dB S/N	/b, f, p/ /ch, d, sh/ /l, w, j/	1.13	/b, f, p/ /ch, d, sh/ /l, w, j/	1.074
−3 dB S/N	/b, ch, d, f, p, sh/ /l, j/ /w/	0.748	/b, ch, d, f, p, sh/ /l, j/ /w/	0.733

the consonants are grouped together for each code as well as the value for $T(x; y)$ in bits generated by the code.

From Table 4 it can be seen that the codings for dyslexics and controls are identical. This suggests that the articulatory features which were best detected were the same for both groups. Moreover, T-test comparisons showed that there was no significant difference in $T(x, y)$ between dyslexics and controls at each masking level (19 dB S/N: $T = 0.75$, $p > .1$. 0 dB S/N: $T = 0.95$, $p > .1$. −2 dB S/N: $T = 1.04$, $p > .1$. −3 dB S/N: $T = 0.53$, $p > .1$). The simplest interpretation of the three-part codes at 19, 0 and −2 dB S/N is as a combination of sonority and placing (i.e., labial vs. non-labial). At maximum masking levels, the three-part codes can best be explained in terms of sonority alone.

5. Discussion

The novel approach used in this experiment was to apply the information analysis of spoken CV confusions to the study of developmental dyslexia. As would be expected from Miller and Nicely's data (1955), the extent of the confusions made by all our subjects increased with increasing noise masking. We found that the pattern of confusions that dyslexics made with our nine CV stimuli was remarkably similar to that made by the controls. In addition, the information analysis showed that the codes which optimized transmitted information were the same for dyslexics and controls at each masking level (a combination of place and sonority). This suggests that at each masking level the two groups were using essentially the same cues to solve the discrimination task. The similarity in performance between dyslexics and controls makes it unlikely that adult dyslexics have widespread distortions in their input phonological representations for these CV sounds. However, we cannot conclude from these data that dyslexics' input phonological representations are normal in every respect. By choosing a restricted set of CV stimuli, we limited our search to a subset of all possible "onsets" in spoken words. Clearly, a comprehensive examination of dyslexics' input phonological representations would require us to measure confusions between a variety of VC "rimes" as well as more complex structures such as CCVs and VCCs which contain co-articulated consonants. In fact there is good reason to suppose that this would be worthwhile. Using Bradley and Bryant's rhyme detection task (1983), researchers have shown that dyslexics' poorest performance is on the rime aspect of the task (Brunswick & Rippon, 1994; Read, 1989). Clearly further experiments using this perceptual confusion technique would be of value. Furthermore, it is possible that while the discrimination of simple CV sounds was too easy a task for adults, it might be more difficult for children. Therefore the same experiment should be carried out in dyslexic children and controls.

While we did not find widespread perceptual confusions for the nine CVs, our dyslexic adults specifically made more confusions between /ʃa/ & /tʃa/ and /fa/ & /pa/, that is, between a fricative and an affricative and between a fricative and a stop. (Note that these are only two out of a total of 72 confusions that were possible with nine stimuli). Closer examination of our stimuli may explain this finding. Fig. 2 shows the results of spectral analysis with a moving FFT window for the stimuli /ʃa/, /tʃa/, /fa/ & /pa/ in the absence of white noise.

The stimulus tʃa is also shown embedded in white noise at the 0 dB S/N ratio. The transition regions in /ʃa/ & /tʃa/ for formants F1, F2 and F3 took about 40–50 ms and were very similar in shape to each other: /ʃa/ = F1 562–625 Hz; F2 1437–1250 Hz; F3 2250–2312 Hz, /tʃa/ = F1 562–625 Hz; F2 1437–1187 Hz; F3 2250–2312 Hz (i.e., *F1 and F2 converged while F2 and F3 diverged*). However, the overall length of the transition region plus vowel was approximately 30–45 ms longer for /tʃa/ than for /ʃa/. The corresponding transition changes for /fa/ & /pa/ were also very similar in shape to each other: they took about 30–40 ms, but were quite different from /ʃa/ & /tʃa/: /fa/ = F1 500–625 Hz; F2 935–1000 Hz; F3 2500–2562 Hz, /pa/ = F1 625–625 Hz; F2 1000–1000 Hz; F3 2562 Hz (i.e., *F1, F2 and F3 all had very shallow, parallel slopes*). Again, the length of the transition region plus vowel was 30–40 ms longer for /pa/ than /fa/. The spectral plot at the bottom of Fig. 4 shows /tʃa/ at 0 dB S/N. It illustrates how, with increasing masking, the plosive components of the stops, the frication components of the af/fricatives as well as F3 and F4 were likely to be below audible threshold (i.e., a signal to noise ratio of −4 dB within one auditory filter, Moore, 1989). Therefore, in the presence of white noise, most of the remaining information to discriminate between /ʃa/ & /tʃa/ and /fa/ & /pa/ is contained within the transition regions plus vowel. However, given that the shapes of the transition regions for /ʃa/ & /tʃa/ and /fa/ & /pa/ were so similar, it is likely that subjects would have had to rely more on the timing differences between these CVs than the spectral cues, in order to tell them apart. If dyslexics have a problem making timing judgements of this kind, it could explain why they made excess confusions between /ʃa/ & /tʃa/ and /fa/ & /pa/. This explanation is consistent with the suggestion by Tallal, Miller, and Fitch (1993) that specifically language-impaired (SLI) children and dyslexics may have difficulty in making judgements based on the temporal aspects of spoken language. This claim is based upon experiments in which SLI children had to associate pairs of synthetic speech sounds (i.e., /ba/-/ba/, /ba/-/da/, /da/-/ba/ or /da/-/da/) with appropriate patterns of button presses (i.e., 1-1, 1-2, 2-1 or 2-2). When the transition region between the consonant and the vowel took only 40 ms, SLI children failed even to associate one button with /ba/ and the other with /da/. When the transition period was extended over 80 ms, SLI children performed as well as controls. Consequently, Tallal et al. (1993) have suggested that the problems experienced by SLI children may be due

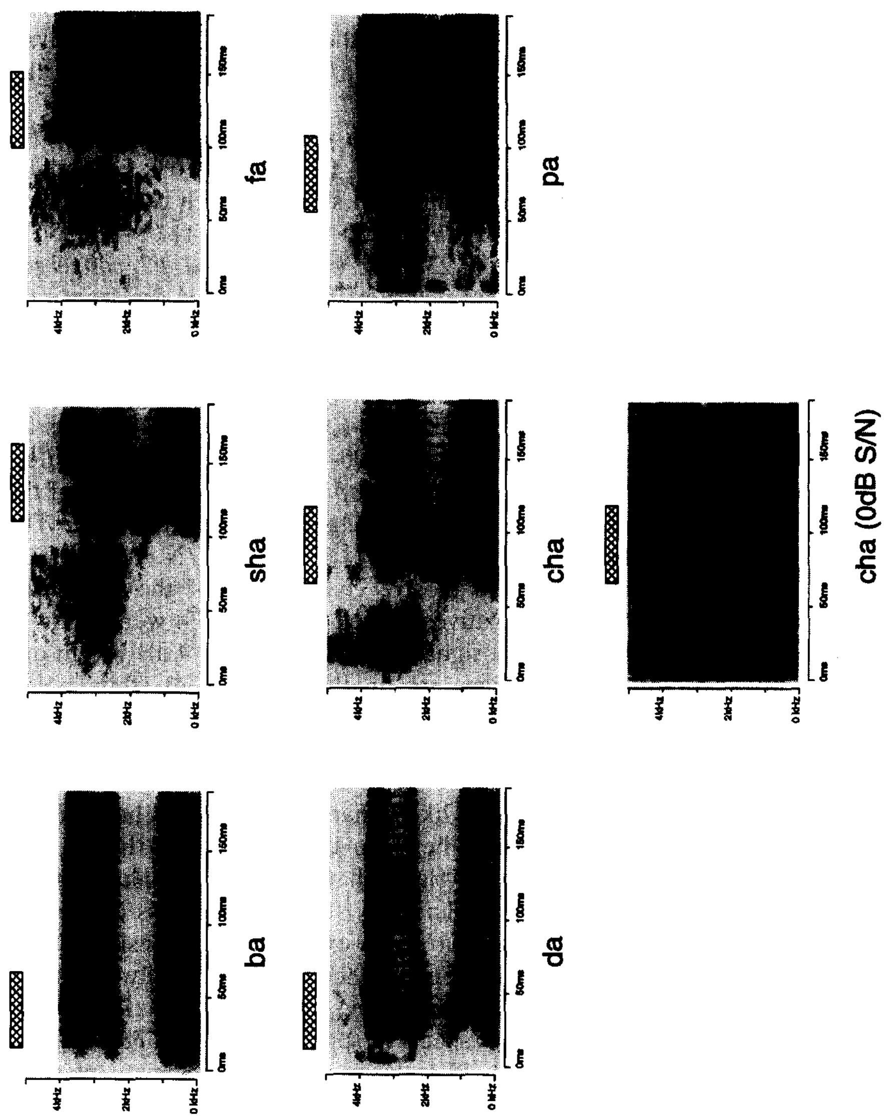


Fig. 2. Spectral plots of /ba/, /da/, /fa/, /pa/, /tʃa/ & /ʃa/ in the absence of white noise mask. The stimulus /tʃa/ is also shown embedded in white noise at the 0 dB S/N level. It shows how the plosive region at the start of the constant, F3 and F4 tend to be masked. In each case, the x-axis represents time (ms), the y-axis frequency (Hz). Darker shading represents increasing amplitude. The cross-hatched bars indicate the location and approximate duration of the transition region between consonant and vowel in each case.

to an inability to resolve speech events which occur within time periods shorter than about 50 ms. This experiment was performed with SLI children, so it is not clear to what extent they can be directly compared with dyslexics. But if we accept Tallal's claim that there are close links between SLI and dyslexia, perhaps our adult dyslexics would have performed like controls if we had exaggerated the timing differences between /ʃa/ & /tʃa/ and /fa/ & /pa/.

EXPERIMENT 2

6. Objective

In Experiment 1, subjects were theoretically able to produce their best possible performance at identifying CVs because they could replay the stimuli in the absence of time constraints. However, it is possible that by allowing subjects to replay stimuli, we may have lost potential differences between the groups: dyslexics may have compensated their difficulty in perceiving some CVs by replaying stimuli several times. To discount this possibility, in Experiment 2 we imposed time constraints by presenting each stimulus once only and recording response times.

7. Methods

The same subjects from the first experiment took part in Experiment 2. The design and protocol were similar to Experiment 1 with two exceptions. First, stimuli were presented either in the absence of noise or at the 0 dB S/N ratio. Secondly, subjects were only allowed to hear the stimuli once per trial. In addition their responses were timed by the computer. The clock was started as soon as the speech stimulus stopped. Timing was stopped at the moment that subjects clicked the mouse button on the response box of their choice.

The instructions given to each subject were as follows: "In this task you will hear a warning beep. Half a second later you will hear one of 9 sounds – ba, cha, da, fa, la, pa, sha, wa or ya. The sound is only played once. You must then move the cursor with the mouse to the appropriate response box as quickly as possible and press the left mouse button. If your choice was correct, the box will light up green. If it was incorrect, the response you should have made will light up red. Remember, you must try to do this task as quickly and as accurately as you can."

Table 5
Experiment 2: confusion matrices

normal

no masking

	b	p	d	f	ch	sh	w	l	y
b	297	0	0	0	0	0	2	0	1
p	1	294	0	0	3	0	0	0	2
d	0	2	296	0	0	0	1	0	1
f	0	2	0	295	0	0	0	2	1
ch	0	0	0	0	297	1	0	1	1
sh	0	0	0	2	19	278	1	0	0
w	0	0	0	0	0	0	296	0	4
l	0	0	0	0	0	0	0	298	2
y	0	0	0	0	1	0	1	0	298
	298	298	296	297	320	279	301	301	310

0dB S/N

	b	p	d	f	ch	sh	w	l	y
b	139	21	6	125	7	2	0	0	0
p	10	260	2	14	8	2	0	1	1
d	0	0	95	0	130	74	0	0	1
f	7	101	3	169	6	13	0	0	1
ch	5	41	38	13	137	64	0	2	0
sh	1	1	1	22	61	214	0	0	0
w	0	1	0	1	2	0	293	1	2
l	1	0	6	0	1	0	4	149	139
y	0	1	0	0	0	0	1	6	292
	163	426	151	344	352	369	298	159	438

dyslexic

no masking

	b	p	d	f	ch	sh	w	l	y
b	293	2	2	0	1	0	0	1	1
p	1	297	0	0	0	1	0	0	1
d	2	2	295	0	0	0	1	0	0
f	0	2	0	297	0	0	0	1	0
ch	0	1	0	0	296	2	1	0	0
sh	0	0	0	1	54	245	0	0	0
w	0	0	1	0	0	0	297	1	1
l	2	0	0	1	0	1	0	296	0
y	0	0	0	0	1	0	4	1	294
	298	304	298	299	352	249	303	300	297

0dB S/N

	b	p	d	f	ch	sh	w	l	y
b	89	45	4	151	4	3	0	3	1
p	7	269	0	8	10	6	0	0	0
d	4	1	63	1	143	85	0	1	2
f	12	158	4	108	6	9	1	2	0
ch	3	34	19	21	152	65	3	0	3
sh	9	5	14	20	114	128	3	3	4
w	1	1	0	2	0	1	291	1	3
l	2	0	4	4	3	0	2	133	154
y	2	1	0	4	1	1	2	4	285
	129	514	108	319	431	298	302	147	452

8. Results

Table 5 shows the raw confusion matrices from Experiment 2. They are displayed in exactly the same way as Table 2 from Experiment 1. As before, subjects made more confusions in the presence of white noise masking. Inspection of the confusion patterns at 0 dB S/N in Tables 2 and 5 shows that subjects' performance in Experiment 2 was very similar to that in Experiment 1. We confirmed this impression with a repeated-measures three-way ANOVA (i.e., Confusions (72 levels) × Group (2 levels) × Experiment (2 levels)). The main effect of *confusion* was significant, $F(71, 1278) = 69.21$, $p < .0005$, while the main effects of *experiment* and *group* were not, $F(1, 18) = 0.69$, $p > .1$; $F(1, 18) = 3.49$, $p > .5$). Furthermore, the interaction *Experiment* × *Group* was not significant, $F(1, 18) = 1.04$, $p > .05$). Thus, both dyslexics and controls produced essentially similar performance whether they were being timed or not. This suggests that subjects' perceptual responses in Experiments 1 and 2 were not significantly affected by the method of stimulus presentation.

Table 6 shows the five confusions which best discriminated between dyslexics and controls at each masking level. Bonferroni multiple T-test comparisons ($\sigma = 0.05$) showed that, in the absence of masking, dyslexics made significantly more ʃa/tʃa confusions than controls. At the 0 dB S/N

Table 6
Experiment 2: the five CV confusions at each masking level which best discriminate between dyslexics and controls

	Confusion	Dyslexics max. = 300	Controls max. = 300	Dyslexics–controls
No mask	sha/cha	54	19	35*
	ja/wa	4	1	3
	wa/ya	1	4	−3
	pa/cha	0	3	−3
	la/ja	0	2	−2
0 dB S/N	fa/pa	158	101	57*
	sha/cha	114	61	53*
	ba/fa	151	125	26
	ba/pa	45	21	24
	cha/da	19	38	−19

* $p < .05$.

masking level fa/pa and ʃa/tʃa confusions best discriminated between dyslexics and controls, although only the former was significant. Thus, as far as the differences between dyslexics and controls are concerned, we obtained similar results to Experiment 1.

Fig. 3 shows the mean times that dyslexics and controls took to respond to each of the nine CV stimuli in the "no mask" and 0 dB S/N conditions. It is clear from Fig. 2 that dyslexics took longer than controls to make their responses. Furthermore, both groups took longer to respond when the white noise mask was present. (Means for "no mask" condition: dyslexics, 1416 ms (*SE* 10.6); controls, 1134 ms (*SE* 7.7). Means for 0 dB mask condition: dyslexics 1533 ms (*SE* 11.1); controls, 1325 ms (*SE* 8.8).

These results were confirmed by a three-way repeated-measures ANOVA of subjects' response times (Group (2 levels) × Stimulus (9 levels) × Masking level (2 levels)). The main effects of *group*, *stimulus* and *mask* were all significant, $F(1, 18) = 18.8$, $p = .004$, $F(8, 144) = 8.9$, $p < .0001$ and $F(1, 18) = 39.9$, $p < .0001$ respectively. All two- and three-way interactions were not significant at $p < .05$ with the exception of *Stimulus* × *Mask*, $F(8, 144) = 7.9$, $p < .0001$. The absence of a significant interaction between group and stimulus suggests that dyslexics were uniformly slower than normals at discriminating spoken CV stimuli.

9. Discussion

In Experiment 1 it is possible that dyslexics could have compensated for a difficulty in perceiving some CV stimuli by replacing them. However, in Experiment 2 we imposed time constraints by presenting each stimulus once only and emphasizing that subjects had to respond as quickly as possible.

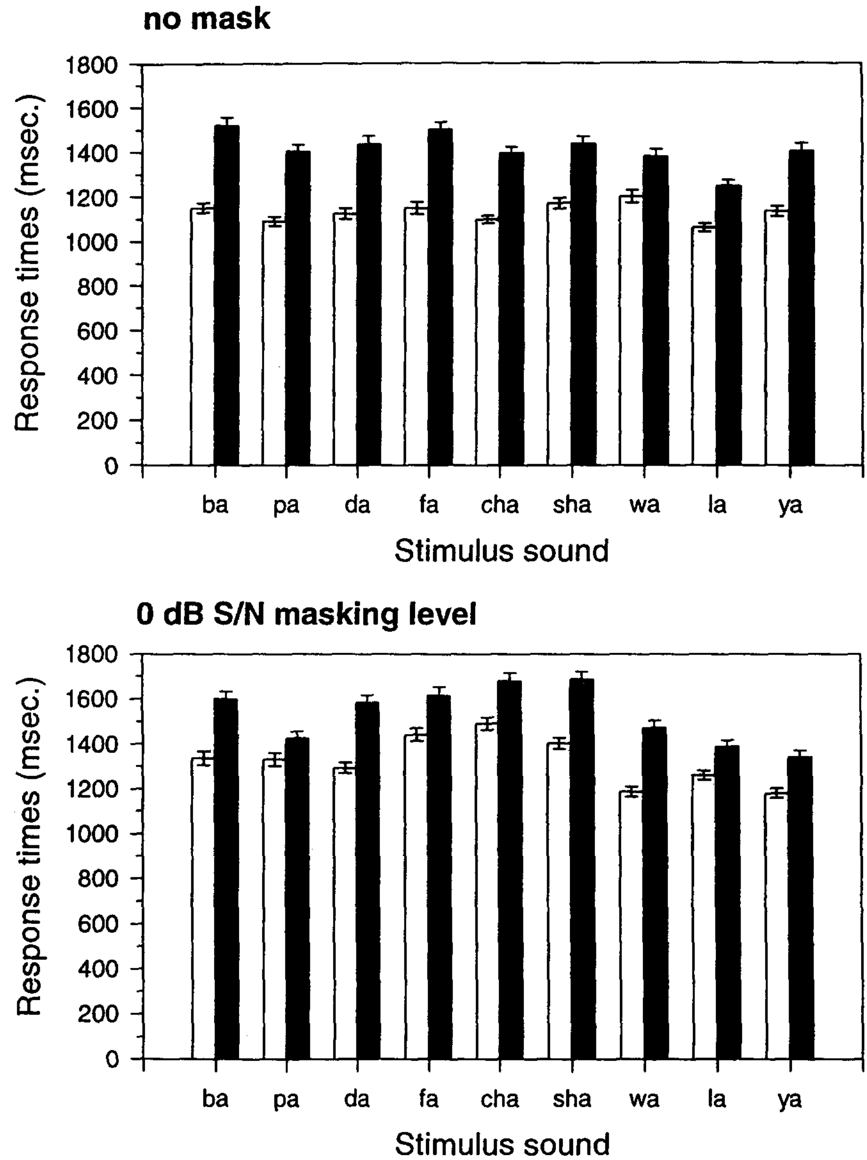


Fig. 3. Bar charts showing the average time that dyslexics (dark bars) and controls (light bars) took to make a response to each of the nine spoken CV stimuli in Experiment 2. The upper graph is for the no mask condition and the lower graph for the 0 dB S/N masking condition. Error bars represent 1 standard error of the mean.

Despite these restrictions, we could not statistically discriminate between subjects' performance in Experiments 1 and 2 at the 0 dB S/N level. In both experiments, the global pattern of performance comparing dyslexics with controls was very similar. But in Experiment 2, we also repeated the finding that dyslexics tended to confuse /pa/ for /fa/ and /tʃa/ for /ʃa/ more than controls. In conclusion, Experiment 2 confirmed that the findings from Experiment 1 were unlikely to have been influenced by the fact that subjects could replay stimuli at their leisure.

The fact that dyslexics' responses were systematically slower than controls requires further explanation. This result could have arisen either because dyslexics were slower at speech processing and/or because they were slower at reading/identifying the written letter strings in the response boxes. If the latter explanation were true, we would expect dyslexics to be equally slow at responding when the stimulus is presented visually. However, if their slow performance reflects slow speech processing, then the difference in response times between dyslexics and controls should be removed by presenting stimuli visually. In Experiment 3, we tested this possibility directly.

EXPERIMENT 3

10. Objective

All subjects performed two versions of a visual letter search task. In the first version of this task, the stimulus was always one of the following visually presented letter strings (ba, cha, da, fa, la, pa, sha, wa, ya). Theoretically, subjects could perform this task in one of two ways. They could either make use of their visual memory for the letter strings, or, by reading the stimulus to themselves, they could utilize the phonological trace thereby evoked to help remember the stimulus. Therefore, we ran a second version of the task to minimize the use of a phonological mediation strategy. In this second version of the task, subjects were presented with unpronounceable letter pairs (cn, gk, jl, mh, tq, vf, wb, xd, zr).

11. Methods

At the beginning of each trial subjects were presented with a central fixation cross. Subjects initiated each trial by pressing a mouse button causing a randomly selected letter string to appear. The centre of the letter string appeared 2° away, NW, NE, SE or SW from the fixation spot. Stimulus letters were the same size as those in the response boxes for the speech experiments and were viewed from the same distance. Each stimulus was presented on screen for 800 ms. 100 ms after the stimulus disappeared, the same nine response boxes as in the speech experiments appeared. Unlike the speech experiments, the locations of the response letter strings were randomized in every trial. Each of the nine stimulus letter strings was presented 20 times, giving a total of 180 trials for each version of the visual search task. Subjects' responses were timed in the same way as Experiment 2.

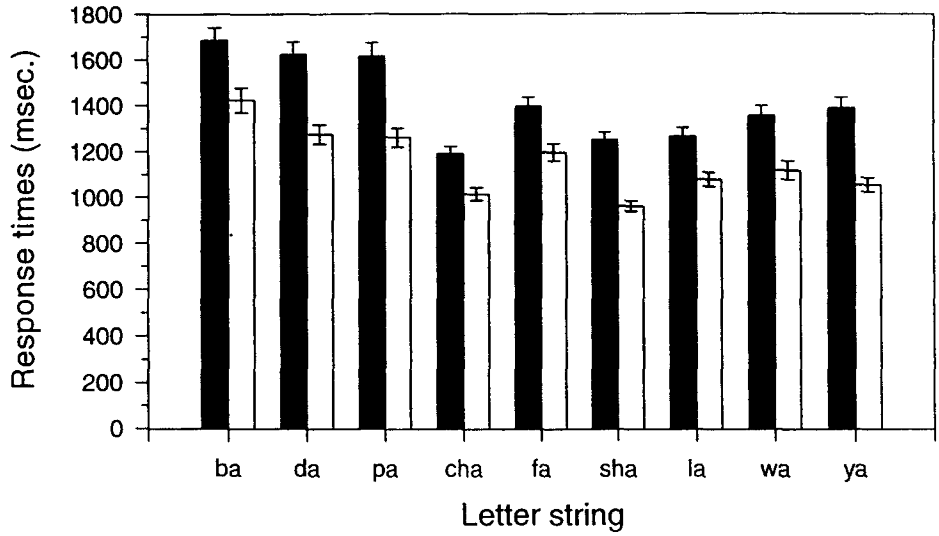
12. Results

Almost all subjects obtained 100% correct responses in both versions of the visual search task and were at ceiling, so we ignored the error data. We assumed that if subjects found some letter strings easier to find than others than this would be reflected in differences in the search times. Fig. 4 shows bar charts of the mean search times that dyslexics and controls took to find the letter strings in the two visual search tasks. It is clear that, overall, dyslexics took longer than controls in both tasks. In the first version of the visual search task, both dyslexics and controls were on average quicker at finding "cha" and "sha" than any of the other letter strings. In the second version of the task, there was an advantage for "cn", "jl" and "mh" compared with the other six letter pairs, but again this pattern was true for dyslexics and controls.

The average response times for dyslexics in condition 1 (pronounceable letter strings) and condition 2 (unpronounceable letter strings) were 1420 ms (*SD* 278) and 1433 ms (*SD* 314) respectively. The equivalent results for controls were 1156 ms (*SD* 250) and 1146 ms (*SD* 261) respectively. We carried out two two-way repeated-measures ANOVAs on the search times, one for each version of the letter search task (Letter string (9 levels) × Group (2 levels)). In both cases, the main effects of *letter string* and *group* were significant (condition 1: $F(8, 128) = 26.6$, $p < .0001$ and $F(1, 16) = 8.9$, $p = .009$; condition 2: $F(8, 128) = 39.5$, $p < .0001$ and $F(1, 16) = 8.7$, $p = .009$). The interactions between *Letter string* × *Group* were not significant (condition 1: $F(8, 128) = 1.3$, $p = .3$; condition 2: $F(8, 128) = 1.6$, $p = .1$). These findings confirm that though some letter strings were easier to find than others, this was true for both groups.

13. Discussion

The difference in response times between dyslexics and controls was very similar for the first condition in Experiment 3 (visually presented, pronounceable letter strings) and Experiment 2. This suggests that dyslexics were not slowed down by the speech perception part of the task in Experiment 2. The fact that both dyslexics and controls found sha & cha quicker than any of the other letter strings may have reflected the fact that they were the only three-letter stimuli; the rest were two-letter stimuli. Together, the results from Experiments 1, 2 and the first part of Experiment



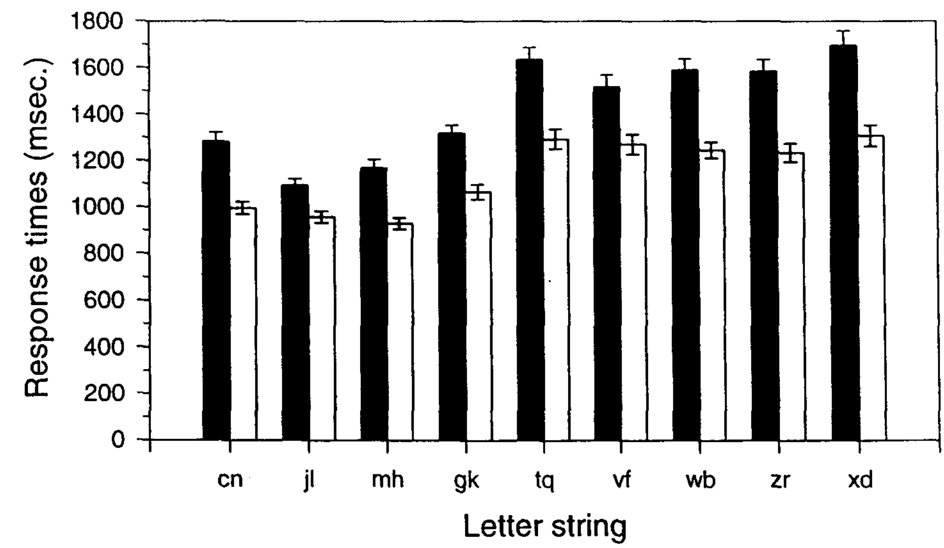


Fig. 4. Bar charts showing the average time that dyslexics (dark bars) and controls (light bars) took to make a response to the visually presented letter strings in Experiment 3. The upper graph is for condition 1 (pronounceable letter strings) and the lower graph for condition 2 (unpronounceable letter strings). Error bars represent 1 standard error of the mean.

3 confirm that dyslexics and controls have very similar auditory perception of CV sounds, with the exception of fa/pa and ʃa/tʃa.

The absence of a timing difference between the two conditions in

Experiment 3 (visually presented, pronounceable letter strings vs. visually presented, unpronounceable letter strings) suggests that the controls did not use a phonological component which was unavailable to the dyslexics. Therefore, either the controls made no use of phonological mediation, or the dyslexics had equal access to this phonological representation. In either case the slower search times of dyslexics in Experiment 2 are unlikely to be accounted for merely by phonological factors. An alternative explanation for these timing differences might be related to the visual perceptual aspects of the tasks, especially in Experiment 3. Not only did subjects have to hold the stimuli in memory, but they also had to find them from the response display. Consequently, the dyslexics might either have been inefficient at committing and retrieving items from visual memory and/or slower than controls at visual search. There is certainly evidence that dyslexics are slower than controls when performing visual search for simple geometric targets (Ruddock, 1991). Like Ruddock, we found a difference in response latency of the order of hundreds of milliseconds between dyslexics and controls which is consistent with impaired visual search. Nevertheless, we also wished to exclude the possibility that our dyslexics had poorer visual memories than controls.

EXPERIMENT 4

14. Objective

In Experiment 4, we compared visual pattern memory in dyslexics and controls.

15. Methods

Our visual memory task was the same as Experiment 3 except that we used visual patterns very similar to Ravens Progressive Matrices instead of letter strings. We designed 20 sets of nine visually similar patterns. One of the nine patterns in each set was presented in isolation as the target. Then all nine patterns appeared on screen, one in each of nine boxes. We measured the time it took for subjects to choose a pattern from the response window and we also recorded whether their response was correct.

16. Results and discussion

Dyslexics and controls correctly identified 76% and 72% of the target patterns respectively. While these are respectable scores, they are unlikely to represent a ceiling effect. A chi-square comparison of the frequency of correct responses between dyslexics and normals was not significant. However, dyslexics did take significantly longer per trial than controls to respond (mean times over 20 trials: 2.6 s (*SD* 0.55) and 2.1 s (*SD* 1.39); $t(18) = 2.15$, $p = .04$). Assuming that we had designed a sensitive task, these results suggest that poor visual memory did not contribute to dyslexics' slow responses in Experiments 2 and 3.

Another contributory factor to account for dyslexics' slowness could be that they were less efficient than controls at operating the computer's mouse interface. Therefore we ran a final experiment to ensure that dyslexics were not systematically less accurate or slower than controls at using the mouse.

EXPERIMENT 5

17. Objective

In this task we wanted to see how quickly and accurately subjects could control the mouse.

18. Methods

Once the subject initiated the experiment, a central fixation point appeared. Simultaneously a target cross appeared randomly at one of eight possible locations which were arranged in a regular ring around the central fixation point. Subjects had to move a square cursor as quickly as possible from the centre of the screen to the target cross where they captured it by pressing the left mouse button. As soon as they had captured the first target cross, a new target appeared randomly in one of the eight locations and the square cursor returned to the centre of the screen. The subject then had to make another rapid mouse movement to capture the next cross. This sequence was repeated until subjects had captured a total of 80 target crosses. They were instructed to complete the task as quickly and as accurately as possible. We measured both how long it took subjects to capture each stimulus as well as the RMS error of their final cursor positions (in screen pixels) with respect to the centre of the target cross. The screen display was VGA resolution (640×480 pixels).

19. Results and discussion

We compared the time it took dyslexics and controls to capture the target crosses as well as the mean RMS error for final cursor position. *T*-tests showed no significant group difference on either measure (mean times 1.4 s (SD 0.23) and 1.25 s (SD 0.14); $t(18) = 1.69$, $p = .1$; mean RMS position 5.3 pixels (SD 1.5) and 5.9 pixels (SD 1.6); $t(18) = -0.81$, $p = .4$). These results suggest that poor mouse control is unlikely to have contributed to dyslexics' slower performance in Experiments 2, 3 or 4.

In conclusion, while we can be sure that our dyslexics' slowness was not specifically linked to speech processing, its cause remains unclear. However, it is plausible that the dyslexics were simply slower at visual search (cf. Ruddock, 1991). The only other possible explanation relates to the strategies that our dyslexics used when carrying out the tasks, rather than any differences in processing time per se. Specifically, our dyslexic subjects knew that they were being tested because they were dyslexic, and expected to be bad in some if not all of the tests that were administered. Therefore they may have been slower because they tried hard to do well and not make mistakes by going too fast.

GENERAL DISCUSSION

The majority of children who can be labelled as developmentally dyslexic perform poorly on the tasks which assess phonological skills. Moreover, most of these children have difficulties reading nonsense words. These two pieces of evidence are usually explained as follows. Since nonsense words cannot be looked up in a mental lexicon, they have to be decoded and their pronunciation assembled by linking orthographic units to phonological segments. One reason that dyslexics perform poorly on phonological tasks might be because the "input" representations of phonological segments are somehow noisy or distorted. If this were true, it is clear that dyslexics would find mapping orthographic units onto phonological units particularly difficult when they tried to read nonsense words.

In this paper, we used a new method to test the hypothesis that dyslexics' input phonological representations might be distorted. We measured the pattern of confusions that dyslexics made between nine spoken CVs and compared them with age matched controls. We assumed that when subjects are asked to classify a spoken stimulus, they must at some level compare what they hear with an internal representation of what each of the nine CV stimuli would be expected to sound like. If dyslexics' representations are different from those of controls, the pattern of responses they give should also be different. We found that the overall pattern of confusions made by the two groups of subjects was very similar for our limited stimulus set. Therefore we concluded that there were no marked, widespread distortions in dyslexics' input representations for these nine CVs. However, of the 72 possible confusions which subjects could have made, dyslexics did make an excess of two of them, confusing /pa/ for /fa/ and /tʃa/ for /ʃa/. This result suggests that while the overall shape of dyslexics' phonological representations for these nine CV sounds is normal, there might be localized distortions within it. A speculative explanation for this difference is that dyslexics might find it harder than controls to discriminate between these sounds when the only useful cue is the length of the transition region plus vowel. It is interesting to note that Masterson, Hazan, and Wijayatilake (1995) showed that the errors made by two phonological dyslexics in a phoneme discrimination task included evidence of problems with certain fricative contrasts.

We propose that this confusion paradigm could usefully be extended to examine dyslexics' input phonological representations further. However, it is conceivable that after experiments with sounds containing VCs, CCVs and VCCs, we may fail to find any evidence for systematic distortion within dyslexics' input phonological representations. How might one then explain dyslexics' difficulties with applying grapheme–phoneme correspondence rules (i.e., sublexical processing)? One possibility, as described in the Introduction, is that their difficulties might be caused by distortions within output rather than input phonological representations. Alternatively, dyslexics might experience problems at the level of graphemic, rather than phonological analysis. This idea implies that the visual systems of poor readers fail to extract invariant letter features from text reliably. Recent evidence has shown that poor performance on low level visual tasks frequently co-authors with phonological difficulties in dyslexics (Slaghuis, Lovegrove, & Davidson, 1993). Moreover, the pattern of reading errors made by poor readers with unstable binocular control can be influenced either by changing print size or by asking children to read with one eye instead of two (Cornelissen, Bradley, Fowler, & Stein, 1991, 1992). Williams, May, Solman, and Zhou (1995) have also shown that manipulating the contrast and spatial frequency properties of letter strings has a profound affect on dyslexics' ability to carry out letter search tasks. Together, these findings suggest that at least some children may not elaborate stable graphemic representations because of subtle abnormalities within their visual systems. If this were true, we have to explain how visual problems might affect dyslexic's ability to apply grapheme–phoneme correspondence rules (the sublexical route) more than their whole word reading (the lexical route). In other words, how might one explain poor nonsense word reading in terms of a visual deficit?

A speculative explanation relies on the different demands that lexical and sublexical processing are likely to make on the system responsible for extracting orthographic features from text. The sublexical route can probably utilize a variety of sizes of orthographic units, while the lexical route is likely to depend on larger units only. Smaller orthographic units, represented by short codes, will be less robust to noise (Cover & Thomas, 1991) whereas larger graphemic units, represented by longer codes, will be more robust to noise. Therefore there should be considerable opportunity for confusion between small graphemic units in the presence of noisy input from the visual system, and hence a greater impact on sublexical than lexical processing. An analogy might be to think of the problem in terms of constructing an identikit face from a set of blurred facial features. If the person reconstructing the face image is only given two eyes and a nose, then there are a large set of real faces with which the impoverished identikit picture might be confused. But if the person is provided with two eyes, a nose, a mouth and so on, the potential for confusion is less.

Finally, it is probably unrealistic to assume that dyslexics' difficulties with grapheme–phoneme processing can be explained in terms of only one of three possibilities: impaired input phonology, impaired output phonology or impaired grapheme processing. There is an abundance of research to suggest that developmental dyslexics represent a heterogeneous population (Boder, 1973; Satz, Morris, & Fletcher, 1985; Watson & Willows, 1993; Seymour, 1986; Castles & Coltheart, 1993). Nevertheless, the majority of dyslexics perform poorly on tests which assess phonological awareness, and it is clear from the studies described earlier that many dyslexics show problems with input phonological representations. Given the evidence presented in this paper, our confusion paradigm provides a powerful method for elucidating the precise nature of these impairments.

Acknowledgements

We would like to thank Professor Burton Rosner of Oxford University Phonetics Laboratory for his advice and both referees for helpful comments on early drafts of this paper. This work was supported by The Wellcome Trust and the McDowell-Pew Centre for Cognitive Neuroscience.

Appendix A. Description of phonetic features

Voicing: The vocal chords vibrate from the outset when the consonants /bdlwy/ are produced, but not for /pftʃʃ/.

Nasality: To articulate /m/ and /n/ the lips are closed and the pressure is released through the nose by lowering the soft palate. The nasal resonance introduced in this way provides an acoustic difference.

Frication: When the articulators close completely, the consonant may be a stop or a nasal. If they are brought close together and air is forced between them, the result is a turbulent noise, frication, characterized by power at high frequencies in the acoustic spectrum /ftʃʃ/.

Placing: This feature describes where in the mouth the major constriction of the vocal passage occurs. A simple, yet comprehensive description of place might distinguish three positions: labial/ bpfw/, alveolar /dl/ and palatal /tʃʃj/.

Sonority: This feature separates sonorants from obstruents. For obstruents the airstream is either fully or partially obstructed during production which is not the case for sonorants. Sonorants comprise the nasals (/mnŋ/) and liquids (lrwj). Obstruents comprise non-nasal stops (/bdkgpt/), fricatives and affricates (/fvszθʒtʃdʒʃ/).

Appendix B. Calculation of transmitted information *T*(*x*; *y*)

Miller and Nicely (1995) decided, a priori, to recode their confusion matrices for a restricted set of phonetic features (voicing, nasality, affrication, duration and place). To illustrate their method using our data, the matrices shown in Table 2 can be recoded for the feature voicing. Accordingly, /bdlwj/ are assigned a "1" for the presence of voicing, and /pftʃʃ/ a "0" for the absence of voicing. At the 19 dB S/N level, control subjects correctly identified 1187 of the non-voiced stimuli and 1480 of the voiced stimuli. However, they misclassified 13 of the non-voiced stimuli as voiced and 20 of the voiced stimuli as non-voiced.

We decided not to make a priori judgements about which feature codes we should use to calculate transmitted information $T(x; y)$. Instead, we calculated $T(x; y)$ for all possible three-part codes for both dyslexics and controls, at each masking level. We then identified which of these codes accounted for the maximum transmitted information. To calculate $T(x; y)$ we used a measure of covariance between input (stimulus) and output (response) (Shannon, 1948). We defined a "mean logarithmic probability" $H(x)$ for the input variable x, which assumes discrete values $i = 1, 2, \ldots, n$ with probability p_i, such that:

$$H(x) = -\sum_i p_i \cdot \log_2 p_i$$

In terms of modern information theory (Cover & Thomas, 1991), $H(x)$ is conventionally termed the entropy of the input variable x, with probability distribution $p(x)$. This can also be thought of as describing the number of binary decisions needed to specify the input (the number of bits of information per stimulus) and is therefore a measure of the uncertainty of the input.

A similar expression holds for the entropy $H(y)$ of the output variable y and the entropy $H(x, y)$ of the stimulus–response pair (x, y). The conditional entropy $H(x \mid y)$ can also be defined, that is, the entropy of the output variable given the input variable. The reduction in entropy, or uncertainty of one variable (output) due to knowledge of the other (input), is given by:

$$T(x;y) = H(x) - H(x \mid y) = H(x) + H(y) - H(xy)$$
$$= \sum_{i,j} p_{i,j} \cdot \log_2 \frac{p_{i,j}}{p_i \cdot p_j}$$

where p_{ij} is the probability of the joint occurrence of input i and output j. This measure, conventionally called the mutual information, is a measure of the covariance of input with output, that is, the transmitted information $T(x;y)$. It is symmetric in x and y and always positive.

References

Annett, M. (1985). *Left, right, hand and brain: the right shift theory*. Hillsdale, NJ: Erlbaum.

Besner, D.E., Twilley, L., McCann, R.S., & Seergobin, K. (1990). On the connection between connectionism and data; are a few words necessary? *Psychological Review*, *97*, 432–446.

Boder, E. (1973). Developmental dyslexia: a diagnostic approach based on three atypical reading–spelling patterns. *Developmental Medicine and Child Neurology*, *15*, 663–687.

Bradley, L., & Bryant, P.E. (1983). Categorizing sounds and learning to read: a causal connection. *Nature*, *301*, 419–421.

Bradley, L.P., & Bryant, P.E. (1985). Rhyme and reason in reading and spelling. *IARLD monographs no. 1*. Ann arbor: University of Michigan Press.

Brandt, J., & Rosen, J.J. (1980). Auditory phonemic perception in dyslexia: categorical identification and discrimination of stop consonants. *Brain and Language*, *9*, 324–337.

Brunswick, N., & Rippon, G. (1994). Auditory even-related potentials, dichotic listening performance and handedness as indices of lateralisation in dyslexic and normal readers. *International Journal of Psychophysiology*, *18*(3), 265–275.

Castles, A., & Coltheart, M. (1993) Varieties of developmental dyslexia. *Cognition*, *47*, 149–180.

Cataldo, S., & Ellis, N. (1988). Interactions in the development of spelling, reading and phonological skills. *Journal of Research in Reading*, *11*, 86–109.

Cornelissen, P.L., Bradley, L., Fowler, M.S., & Stein J.F. (1991). What children see affects how they read. *Developmental Medicine and Child Neurology*, *33*, 755–762.

Cornelissen, P.L., Bradley, L., Fowler, M.S., & Stein J.F. (1992). Covering one eye affects how some children read. *Developmental Medicine and Child Neurology*, *34*, 296–303.

Cover, T.M., & Thomas, J.A. (1991). *Elements of information theory*. New York: Wiley-Interscience.

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

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. London: Erlbaum.

Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, *7*, 6–10.

Jorm, A.F. (1983). Specific reading retardation and working memory: a review. *British Journal of Psychology*, *74*, 311–342.

Kochnower, J., Richardson, E., & DiBenedetto, B. (1983). A comparison of the phonic decoding ability of normal and learning disabled children. *Journal of Learning Disabilities*, *16*, 348–351.

Manis, F., McBride, C., Seidenberg, M., Doi, L., & Custodio, R. (1993). *Speech perception and phonological awareness in normal and disabled readers*. Poster presentation at meeting of the SRCD, New Orleans.

Masterson, J., Hazan, V., & Wijayatilake, L. (1995). Phonemic processing problems in developmental phonological dyslexia. *Cognitive Neuropsychology 12*, 233–259.

Miller, G.A., & Nicely, P.E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*(2), 338–352.

Moore, B.C.J. (1989). *An introduction to the psychology of hearing* (3rd edn). London: Academic Press.

Olson, R.K. (1985). Disabled reading processes and cognitive profiles. In D. Gray & J. Kavanagh (Eds.), *Biobehavioural measures of dyslexia*. Parkton, MD: New York Press.

Patterson, K.E., Marshall, J.C., & Coltheart, M. (1985). *Surface dyslexia*. London: Academic Press.

Pralong, D., & Carlisle, S. (1994). Measuring the human head-related transfer functions: a novel method for the construction and calibration of a miniature in-ear recording system. *Journal of the Acoustical Society of America, 95*(6), 3435–3444.

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

Ruddock, K. (1991) Visual search in dyslexia. In Stein, J.F. (Ed.), *Vision and visual dyslexia: Vol. 13 of vision and visual dysfunction*. London: Macmillan Press.

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal of Child Psychology and Psychiatry, 16*, 181–197.

Satz, P., Morris, R., & Fletcher, J.M. (1985). Hypotheses, sub-types and individual differences in dyslexia: some reflection. In D.B. Gray & J.F. Kavanagh (Eds.), *Biobehavioural measures of dyslexia*. Parkton, MD: New York Press.

Schonell, F.J. (1950). *Diagnostic and attainment testing*. London: Oliver & Boyd.

Seymour, P.H.K. (1986). *Cognitive analysis of dyslexia*. London: Routledge & Kegan Paul.

Shallice, T., & Warrington, E.K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K.E. Patterson & J.C. Marshall (Eds.), *Deep dyslexia* (pp. 119–145). London: Routledge & Kegan Paul.

Shankweiler, D., & Liberman, I.Y. (1972). Misreading: a search for a cause. In J.F. Kavanaugh & J. Mattingly (Eds.), *Language by ear and by eye*. Cambridge, MA: MIT Press.

Slaghuis, W.L., Lovegrove, W.J., & Davidson, J.A. (1993). Visual and language processing deficits are concurrent in dyslexia. *Cortex, 29*, 601–615.

Snedecor, G.W., & Cochran, W.G. (1967). *Statistical methods* (6th edn.), Ames, IA: Iowa State University Press.

Snowling, M.J. (1980). The development of grapheme–phoneme correspondences in normal and dyslexic readers. *Journal of Experimental Child Psychology, 29*, 294–305.

Snowling, M.J. (1991). Developmental reading disorders. *Journal of Child Psychology and Psychiatry, 32*, 49–77.

Snowling, M.J., Hulme, C., Smith, A., & Thomas, J. (1994). The effects of phonetic similarity and list length on children's sound categorization performance. *Journal of Experimental Child Psychology, 58*, 160–180.

Snowling, M.J., & Rack, J. (1991) Dyslexia: defects in the development of grapheme–phoneme correspondence? In Stein J.F. (Ed.), *Vision and visual dyslexia: Vol. 13 of vision and visual dysfunction*. London: Macmillan Press.

Snowling, M.J., Stackhouse, J., & Rack, J. (1986). Phonological dyslexia and dysgraphia: a developmental analysis. *Cognitive Neuropsychology, 3*(3), 309–339.

Stanovich, K.E. (1991). The theoretical and practical consequences of discrepancy definitions of dyslexia. In M.J. Snowling & M.E. Thomson (Eds.), *Dyslexia: integrating theory and practice*. London: Whurr.

Steffens, M.L., Eilers, R.E., Gross-Glenn, K., & Jallad, B. (1992). Speech perception in adult subjects with familial dyslexia. *Journal of Speech and Hearing Research*, 35, 192–200.

Tallal, P., Miller, S., & Fitch, R.H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. In P. Tallal, A.M. Galaburda, R.R. Llinas, & C. von Euler (Eds.), *Temporal information processing in the nervous system. Annals of the New York Academy of Sciences, 682*.

Treiman, R., & Zukowski, A. (1991). Levels of phonological awareness. In S.A. Brady & D.P. Shankweiler (Eds.), *Phonological processes in literacy*. London: Erlbaum.

Tunmer, W.E. (1989). The role of language-related factors in reading disability. In D. Shankweiler & I.Y. Liberman (Eds.), *Phonology and reading disability*. Ann Arbor: University of Michigan Press.

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

Watson, C., & Willows, D.M. (1993) Evidence for a visual processing deficit subtype among disabled readers. In D.M. Willows, R.S. Kruk & E. Corcos (Eds.), *Visual processes in reading and reading disabilities*. Hillsdale NJ: Erlbaum.

Werker, J.F., & Tees, R.C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology*, *41*(1), 48–61.

Williams, M.C., May, J.G., Solman, R., & Zhou, H. (1995). The effects of spatial filtering and contrast reduction on visual search times in good and poor readers. *Vision Research*, *35*(2), 285–291.