Scandinavian Journal of Psychology, 2005, 46, 375–384

Health and Disability

Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers

CARSTEN ELBRO and METTE NYGAARD JENSEN

Department of Nordic Studies and Linguistics, University of Copenhagen, Denmark

Elbro, C. & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, *46*, 375–384.

This study of dyslexia was concerned with the quality of phonological representations of lexical items. It extended the studies of verbal learning in dyslexia from learning new vocabulary items (pseudo-names) to the learning of more well-specified variants of known words. The participants were 19 dyslexic adolescents in grades 4 to 6 and 19 younger normal readers in grade 2 matched on single word decoding. The dyslexics were significantly outperformed by the reading-age controls in non-word reading and in phoneme awareness. The dyslexics also took longer time to learn to associate a set of pseudo-names with pictures of persons although the dyslexics learned to associate familiar names with pictures as quickly as the controls did. The acquisition of new phonological representations of words was studied in an imitation task with maximally distinct pronunciations of long, familiar words. The dyslexics gained less than the controls in this task. They also gained less on one measure taken from a phoneme substitution task with the same words as in the distinctness task. The results are interpreted in the light of the hypothesis that poorly specified phonological representations may be an underlying problem in dyslexia.

Key words: Phonological representation, reading difficulties, dyslexia, phoneme awareness, dynamic testing, paired-associate learning.

Carsten Elbro, Department of Nordic Studies and Linguistics, University of Copenhagen, 86 Njalsgade, DK-2300 Copenhagen S, Denmark. E-mail: CE@cphling.dk

INTRODUCTION

A significant proportion of children's reading problems can be explained by poor word decoding, be it either inaccurate, or slow and non-automatic, or both. Such severe problems with the acquisition of decoding skills in reading are generally referred to as dyslexia. Insufficient phonological processing is possibly the most robust finding concerning the linguistic roots of dyslexia (Lyon, 1995). Some problems with phonological processing are particularly severe and continue even into adulthood (Elbro, Nielsen & Petersen, 1994; Pennington, Lefly, Van Orden, Bookman & Smith, 1987). These problems appear to be most pronounced at the level of individual speech sounds, or phonemes. The fact that dyslexics have such problems with segments of spoken language should not come as a surprise. After all, the letters of alphabetic orthographies represent segments of spoken language at the phoneme level. A problem with access to segments of spoken language is, thus, a problem with access to the spoken units that letters represent.

The precise nature of the phonological processing problems in dyslexia is a central topic in recent psycholinguistic research on dyslexia. One suggestion is that the phonological representations of lexical items may be less well specified than normally (Elbro, 1996; Foy & Mann, 2001; Griffith & Snowling, 2002). This is certainly not the only hypothesis about the nature and origin of phonological processing problems in dyslexia. But it is the (general) hypothesis with which the present study was concerned.

Phonological representations are mental representations of the spoken units of language (e.g. words) – stored in long term memory. Such mental representations may contain more or less information about the spoken units – both within a language user and between language users. A poorly specified representation is one that specifies only parts of the phonetic material of the unit. For example, it may reflect only a short form such as "sub" for *subway* or *submarine*, or "croc" for *crocodile*. Or a poorly specified phonological representation may only provide a full specification of some of the segments of the lexical item, e.g. "cro?dile" where the question mark indicates that some segment should be there but not which. The "cro?dile" representation may lead to deviant pronunciations such as "crowdile" or "crowledile" when the unspecified segment is given whatever sound that fits with the phonotax of the language.

In fact, it is quite possible that a phonological representation may be poorly specified without any overt consequence for normal speech production or perception. Consider, for example, the following hypothetical representation of the spoken form of the word *string* (from Elbro, 1996).

/s	t	r	i	ŋ/
−vocal.	+cons.	[+cons.]	+vocal.	+back
+cons.	+ant.		−cons.	+nasal
−high	+coron.		+high	
−back			−back	
−low				
+ant.				
+coron.				
−voice				
+contin.				
−nasal				
+strid.				

The word *string* is uniquely specified by the above representation even though four out of five segments are very under-specified by the features listed, i.e., they could be other sounds than the actual ones in the word *string*. The reason is that contextual (phonotactic) constraints may help define the sound segments. For example, a consonant [+cons.] following initial [st-] can only be an [r].

If poorly specified phonological representations, such as the hypothetical one above, are common in dyslexia, it would help explain the problems with phoneme awareness that are typical in dyslexia. A dyslexic person with the above representation of *string* would not be able to delete the first two segments and say "string" without the initial "st-". He or she would only know that the remaining part of the word began with a consonant, not which consonant. In general, phonemic segmentation of words is a very difficult task if the relevant segments are not specified in the representation of the words (e.g. Elbro, 1996, 1998; Elbro, Borstrøm & Petersen, 1998; Foy & Mann, 2001; Fowler, 1991; Griffith & Snowling, 2002; Swan & Goswami, 1997; Wesseling & Reitsma, 2001). Yet, the dyslexic person would still be perfectly able to identify the word *string*, remember it, and pronounce it. In other words, phoneme awareness would be specifically impaired while vocabulary, language perception and production would be relatively unimpaired.

If the *acquisition* of well specified phonological representations is a problem in dyslexia (see below), it may also provide a more direct explanation of reading difficulties in dyslexia. Rapid written word recognition may depend on the development of spelling-linked *phonological* representations in addition to already acquired, standard representations, e.g. [jɑːʃt] in addition to standard [jɒt] for *yacht*. Such spelling-linked pronunciations would allow for a *direct mapping* of orthographic to phonological representations in both spelling and reading (Mayringer & Wimmer, 2000; Rack, Hulme, Snowling & Wightman, 1994; Wimmer, Mayringer & Landerl, 2000). Figure 1 illustrates this and the above mentioned relationships between quality of phonological representations, phoneme awareness, and reading development.

Previous research provides some evidence that dyslexia may indeed be associated with poorly specified phonological representations of lexical items.

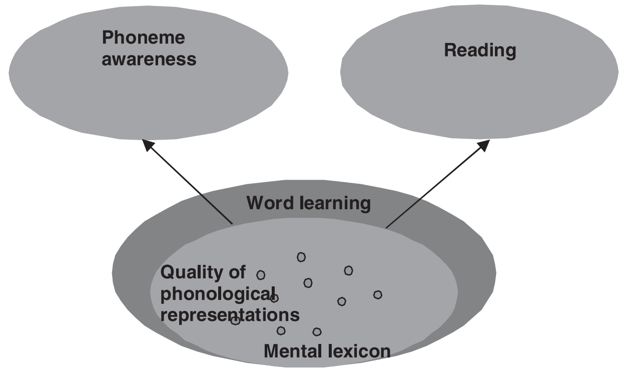

Fig. 1. Hypothetical influences of quality of phonological representations on phoneme awareness and reading. Dots in the mental lexicon represent phonological representations of lexical items.

First, there is some correlational evidence. Adults with a history of reading difficulties have specific receptive vocabulary deficits when words and distractors are phonologically similar. Such adults have also been found to pronounce long, familiar words at a somewhat lower level of distinctness than normal readers do (Elbro, Nielsen & Petersen, 1994). Similarly, young dyslexics have poorer phonological representations of long words than younger reading-age-matched controls (Nation, Marshall & Snowling, 2001). Dyslexics are also less able than normal readers to correct a slightly deviant pronunciation of long familiar words even when receptive vocabulary is controlled (e.g., Fowler & Swainson, 1999).

Second, distinctness of phonological representations have been found to *predict* both development of phonological awareness and reading (Elbro *et al.*, 1998; Wesseling, 1999; Elbro & Petersen, 2004).

Third, quality of phonological representations *predicts response* to phoneme awareness training in the kindergarten grade in children of dyslexic parents (Elbro, 1998; Petersen, 2000).

Fourth, a pilot study indicated that improvements of kindergarten children's phonological representations following training were associated with progress in phoneme awareness with the trained words (Elbro & Pallesen, 2002), but not with untrained words. Although the within-subjects design of this study does not allow for conclusions about causality, it is in concord with the possibility that quality of phonological representations is a determinant of the development of phoneme awareness – at least for specific words.

These results raise several questions, one of which is *why* dyslexia may be associated with less well specified phonological representations of words. Very little work has been done on the possible causes of poorly specified phonological representations. One possible explanation is offered by the lexical restructuring hypothesis. According to this hypothesis, phonological representations are gradually reorganized into ever smaller segments during language acquisition (from 0 to 8 years of age) (Fowler, 1991; Metsala, 1999; Walley, 1993).

Initially, the units of representation are whole utterances, phrases, or words; but as vocabulary increases, the representations are gradually restructured into smaller units – ultimately into phoneme size segments. This development is reflected in the development of phonological awareness: normally children are first aware of words as units of expression, only later do they become aware of smaller segments such as single phonemes. However, the lexical restructuring hypothesis does not readily explain why dyslexics continue to have very poor phoneme awareness even though their vocabulary grows at a near normal rate. For some unknown reason, lexical restructuring seems to take place at a lower rate than in other children.

Another, but far from the last, possibility is that the *acquisition* of phonological representations is generally impeded in dyslexia. A problem with the establishment of well-specified phonological representations could reflect a more general problem with *phonological encoding* (cf. Brady, 1997). This possibility would be in accordance with the well-documented phonological-lexical learning problems in dyslexia. For example, there is rather unanimous evidence that dyslexia is associated with learning difficulties in a paired associate paradigm with *new* verbal material (e.g., Aguiar & Brady, 1991; Johnston & Scott, 2000; Kamhi, Catts & Mauer, 1990; Vellutino, Steger, Harding & Phillips, 1975). For example, Mayringer and Wimmer (2000) reported that German speaking 9-year-old dyslexics had problems learning new words (pseudonames), but performed at the level of their peers in paired associate learning tasks with known, short words.

Parallel results have been reported from studies of *immediate repetition* of words and non-words in dyslexia. For example, Snowling, Goulandris, Bowlby and Howell (1986) found that dyslexics had problems relatively to both chronological age controls and reading age controls with immediate repetition of non-words, while the dyslexics performed at a normal level with repetition of high-frequency real words.

Hence, it may be that dyslexics have a general problem with the acquisition of phonological-lexical material. A general problem with phonological-verbal learning might hypothetically extend into the acquisition of *additional specifications* of already acquired phonological representations of lexical items. The main purpose of the present study was to explore this hypothesis.

Phonological representations are mental entities and can only be studied indirectly. Several techniques have been employed in previous studies, e.g., detection and correction of pronunciation errors, recognition of words based on the first auditory fragment of them (as with the "gating" technique), or recognition of words in noise. In the present study, we employed the "poor parrot" paradigm, in which participants are asked to teach correct pronunciations to a "parrot" (a hand-held toy) with a speech problem. This way we attempted to tap into the most well specified representation of each of the participants' phonological representation. Even pre-school children readily understand this task and interpret it in much the same way when they say the difficult words slowly and as clearly as they can to the "poor parrot" (Elbro *et al.*, 1998).

Just like all other measures of productive language abilities, the poor parrot paradigm is dependent on the participant's ability to pronounce words according to intension. A participant may say "fis" when intending to say *fish*, or "wobin" for *Robin* because of limited articulatory control, in which case the paradigm will provide an unreliable measure of the phonological representation. It is possible to control for articulatory problems. For example, the experimenter may repeat the participant's pronunciation (on behalf of the poor parrot) and ask the participant whether or not the pronunciation is correct. This procedure was adopted in a study with Danish preschool children (Elbro *et al.*, 1998), but the 6-year-old children in the study very rarely rejected their own pronunciation. Furthermore, pronunciation difficulties usually follow identifiable patterns so that it may be inferred whether or not a mispronunciation is likely to be of articulatory or lexical origin.

The poor parrot paradigm is unlikely to work well with short, high frequency words because they have little scope for individual variation. The paradigm probably works best with long, well-known words (or other lexicalized items) which may be pronounced more or less distinctly, i.e., more or less well specified. With children and young adults, words like *chocolate*, *crocodile*, and *library* may be good choices. In a *maximally distinct pronunciation* these words retain four, four and three syllables, respectively, whereas the more frequent pronunciations in coherent speech have fewer segments and syllables, "crocdile", "choclate", and "libri". The maximally distinct pronunciations are often the ones listed in dictionaries even though they may be quite rare and occur only when the words are pronounced carefully in isolation. The optional variation in pronunciation means that there is room for individual variation, also in the way the words are represented. If a person has access to the less well specified everyday representations *only*, this may have few consequences for normal oral communication. It is, therefore, a possibility that the poor parrot paradigm may bring subtle differences to light; and perhaps dyslexia is associated with poor access to the maximally distinct variations of words.

The specific research questions were the following:

(1) Do dyslexics possess less well specified phonological representations than younger normal readers (reading age controls) – even though the older dyslexics may have larger vocabularies?
(2) Do young dyslexics acquire new verbal material (non-words) with more difficulty than younger reading level controls?
(3) Do dyslexics acquire well-specified representations of *known* words with more difficulty than reading age controls when given the opportunity?
(4) Are word-specific gains in quality of phonological representations associated with word-specific gains in phoneme awareness?

The hypotheses were positive replies to each of the research questions.

METHOD

The study employed a reading-level-match design with severely dyslexic adolescents and younger, normally achieving children matched for written word decoding ability. Two measures of verbal learning were administered to both groups, one with new words (pseudonames) and one with maximally distinct pronunciations of well-known, long words. Medium-term retention was measured after one week. The study also looked into the effects of learning maximally distinct pronunciations on phoneme awareness with the *same words* in a within subjects design.

Participants

Nineteen dyslexic adolescents in grades 4 to 6 (12 boys and 7 girls) were selected from two special schools for dyslexics. They had been referred to these schools after failure to learn to read in both ordinary classes and with special instruction. Their level of word decoding corresponded on average to that reached by normal readers who were three and a half years younger. The mean age of the dyslexics was 12 years 1 month (*SD* = 8 months). They scored 46.8 correct (*SD* = 8.4) on a test of single word decoding (see below), a score which is close to average in May in grade 2. Nineteen younger normal children (10 boys and 9 girls, age 8 years 6 months, *SD* = 5 months) were selected from second grade classes in schools in Copenhagen following screening of whole classes. Based on the screening results, children were selected to match the dyslexic group on word decoding (mean score and variance) – avoiding children in normal classes with extreme scores relative to their class mates. This selection procedure meant that the reading scores fell within a relatively limited range (mean = 46.6, *SD* = 11.8). The normal control children scored very close to the expected grade level, which is 47.9 correct (*SD* = 17.6) in May in second grade. Danish was the first language of all participants. None of the participants had been diagnosed with general or specific language delays. No participant showed signs of articulatory problems that might endanger the reliability of productive language measures.

Materials

Word decoding. Participants were matched on silent decoding of real words ("Ordlæs"; Borstrøm, Petersen & Elbro, 1999). Each item of the test had a pictured object and four words from which participants were supposed to choose the correct one. The distractor words share some of the letters with the correct words so as to reduce successful guessing based on single letters. Participants were given 5 minutes to solve as many of the 112 items as possible. The score was the number of correct responses within this time limit. Average scores from more than 3,000 children are available from the beginning of grade 1 until the end of grade 3. Test-retest correlation is 0.89 between May in grade 2 and October in grade 3, which suggests high reliability. The correlations between this and other word decoding tests for this age group are equally high, suggesting that the validity is high. This and the following two measures were administered as group measures to whole classes.

Non-word decoding. Non-word decoding was measured in a silent pseudo-homophone decision task. In each trial, the participants are presented with 4 non-words and asked which one would sound like a real word if read aloud ("Lis skriver ord", Borstrøm *et al.*, 1999). Participants were given five minutes to solve as many of the test's 60 items as possible. The score was the number correct within the time limit. Test-retest correlation is 0.85 between the end of grade 2 and the beginning of grade 3. The test correlates strongly with oral reading of non-words (Elbro *et al.*, 1994).

Receptive vocabulary (OK test form E for children; Grønborg, Lund, Møller & Petersen, 1993). In this standardized vocabulary measure, participants are asked to mark the correct picture out of three from roughly the same semantic field upon hearing the word (similarly to the Peabody Picture Vocabulary Test). The test has 23 items, and the score is number of correct. The test is the only standardized vocabulary test for children aged 7–9. The manual does not provide data on reliability. However, the adult version of the test has been reported to correlate at $r = 0.64$ with another test of receptive vocabulary in a large adult sample (Elbro *et al.*, 1994).

Phoneme awareness. Participants were asked to substitute all vowels with "ee" [i] in each of 27 frequent, polysyllabic words (listed in the appendix). This and the following tasks were administered to the participants individually. Two scores were taken from this task. The first was the proportion of correctly substituted vowels in intact syllables, i.e. vowels were substituted correctly while no other alternations were made. Test-retest correlation in the present study was 0.92. The second score was the proportion of intact syllables, i.e., the consonants remained unaltered – regardless of whether the vowel was substituted or not. This second score was computed because we wished to assess the robustness of the phonological representations while participants were attempting to substitute the vowels. Test-retest correlation was 0.89.

Quality of phonological representations was measured by means of the "poor parrot" paradigm using the same 27 items as in the phoneme awareness task. The "poor parrot" paradigm (Elbro *et al.*, 1998) works as follows: The participant is introduced to a hand-held doll or "poor parrot" and told that the doll has a problem with speaking. The task of the participant is to teach the doll to pronounce words correctly. The experimenter moves the doll and speaks on behalf of it. The participant and the doll are then shown pictured objects which the doll (i.e. experimenter) names in a very reduced form, e.g. "codi" for *crocodile*. The participant is to offer the correct pronunciation for the doll. It is stressed that the participant has to be very careful to pronounce each word as accurately and clearly as possible. There are two trials with each word. All sessions were tape recorded for later analysis. Scoring was done by two trained raters; they agreed about the phonetic transcription of 96% of the responses. Inter-rater agreement was previously reported at 93% in a study of preschool children (Elbro *et al.*, 1998). Two measures of quality of pronunciation were taken from the participant's best pronunciation. The first was a simple proportion of the words that was pronounced *correctly* (i.e., conventionally according to the unabridged dictionary of Danish pronunciation; Brink, Heger, Lund & Jørgensen, 1991). Test-retest correlation was 0.72 across all items, which provides a lower bound estimate of the reliability because some but not all items were trained between test and retest. The second measure was the *level of distinctness* – scored as the percentage of target vowel segments (marked in the appendix) that were pronounced at the maximally distinct level. The target vowels were selected because they can be pronounced more or less distinctly in normal speech according to the pronunciation dictionary (Brink *et al.*, 1991). For example, in *chok<u>o</u>lad<u>e</u>* ("chocolate"), the scoring was only concerned with the pronunciation of the middle *o* and the final *e*. In a fully distinct pronunciation (as specified by the pronunciation dictionary), the target *o* is pronounced [o] and the *e* as [ə] (the weak schwa vowel). This pronunciation would yield a score of 100% (with this word). A pronunciation with a reduced *o*, e.g. to [ə] or nothing, would yield a score of 50%, provided that the final *e* was still pronounced. If both vowels were reduced, the score would be 0%. This scoring is a simplification of the one used in previous studies (e.g., Elbro *et al.*, 1998). It was adopted because of its simplicity and because it matched the goal of the learning task with distinct phonological representations (see below). Test-retest correlation was

0.79 across all items, which is a lower bound estimate of the reliability because some, but not all items were trained between pre- and post-test.

Word and non-word learning. Paired associate learning of words and non-words with pictures was one of two measures of verbal learning. In the paired associate task (modelled after Mayringer & Wimmer, 2000), participants were shown pictures of human faces and asked to associate real names and non-words with them. This task was selected because it has been employed in a number of previous studies of dyslexia and because it provides an easy way of studying both ease of learning and medium-term retention. The pictured faces were quite distinct because of ethnic and gender differences. There were two conditions, one with five real names (*Helle*, *Rune*, *Lone*, *Jeppe*, and *Nina*, all frequent in Danish) and one with four novel names (*Pøne*, *Fitte*, *Lube*, and *Sylla*, all non-words in Danish). Initially, participants were asked to repeat each name after the experimenter in order to make sure that the participant had heard the name (word or non-word) and was able to pronounce it correctly. In each condition, participants were given a maximum of 15 trials to learn to associate names with pictures. Correct naming of all pictures in two consecutive trials was required. The measure was number of trials taken to learn all pictures in each set. The score was set to 16 when a participant failed to learn the word-picture associations in 15 trials. In a follow-up session one week later, the photographs were presented to the participants and they were asked to name them. The measure was number of correctly named pictures. Only the non-word condition required new verbal material to be learned. It was hypothesized that the dyslexic adolescents would have problems in this condition. The word condition served as a control for basic associate learning abilities.

Learning fully distinct phonological representations. This was the second measure of verbal learning. This task focused on words that were already known to the participants. The aim of the measure was to study how easily participants acquired the most distinct pronunciation of these words. The measure employed a simple design with pre-test, training and post-test. At pre- and post-test, the participants' quality of phonological representations was assessed by means of the "poor parrot" measure described above. Training comprised nine of the 27 test words (see the appendix). These words were *trained in an imitation game* in which participants were asked to try to say each word in the same way as a computer-generated voice. The voice was introduced to the participants as the voice of a robot. The target pronunciation was delivered by high quality synthetic speech – at the highest normal level of distinctness (following Brink *et al.*, 1991). Participants were asked to imitate the target pronunciation five times with each word. It should be noticed that this imitation training did not direct the participants' attention to single segments of words, something which is very hard to avoid when instruction is given by humans. Another nine words (of the 27 words) were just *repeated* quickly five times each (following the procedure of Stone and Brady, 1995). The remaining nine words served as *untrained* control words. The words were presented in the same order to all participants. The question was whether the older dyslexics would find it harder than the RA controls to improve the quality of their representations of the trained words.

Procedure

Following the initial group testing of reading and receptive vocabulary, the children were seen individually twice. On the first occasion, the participants' pre-test scores were taken, and the participants were taught distinct phonological representations. This lasted about 45 minutes in all. On the second occasion, one week later, post-test scores were taken, and medium-term word learning was assessed. This took about 20 minutes.

Analyses

The significance of group mean differences was tested by means of *t*-tests. One exception was the word and non-word learning task where failures to learn the word-picture associations in 15 trials were given a (somewhat arbitrary) high score of 16. Hence, scores with this task were analysed by means of non-parametric tests. All individual gains were calculated as differences between log-odds transformed pre- and post-test scores. The log-odds (or *lod*) transformation was preferred because a standard repeated measures ANOVA may yield grossly misleading results when performed on raw accuracy scores outside a narrow interval around 50% correct (see Allerup and Elbro, 1998). Log-odds transformations are standard where life or money are at stake such as in medicine and horse racing. Hence, the traditional time variable (pre- to post-test) was contained in the individual gains scores. The individual gain scores in both the distinctness learning task and in the phoneme awareness task were subjected to a 2 groups (dyslexic vs. RA controls) × 3 conditions (trained, repeated and untrained words) repeated measures ANOVA. The alpha level for this study was 0.05.

RESULTS

Results with the pre-test measures are displayed in Table 1. The two groups differed significantly on the measure of non-word decoding (pseudo-homophone detection) even though they matched closely on silent word decoding. This finding is in accordance with previous research (Rack, Snowling & Olson, 1992) and suggests that a characteristic of dyslexia is a particular weakness with phonological coding in reading. This weakness is likely to reflect basic problems with the acquisition of the alphabetic principle of the orthography.

The dyslexics scored higher than the younger controls on a measure of receptive vocabulary, which is not surprising given that they were about three and a half years older. The difference was not significant, though. As expected, the dyslexics scored much lower than the younger controls in the phoneme awareness task (with phoneme substitution), at pre-test with all 27 words. In addition to making fewer correct substitutions, the dyslexics also often changed *other* parts of the syllabic structure than just the vowels – and they did so more often than the controls did (21.3% vs. 9% of the responses). For example, *helicopter* [helikʌptʌ] might become [hililipti(r)] rather than the correct [hilikipti(r)] (see Table 1).

The dyslexic participants scored lower than the younger normal readers on the measure of quality of phonological representations. The dyslexics produced twice as many errors as the controls did. However, the group difference was not significant in a parametric test, possibly because of a ceiling effect in the normal control group. Since the scores were not normally distributed (Kolmogorov-Smirnov test statistic 0.32, $p < 0.001$), a non-parametric analysis of group means was conducted. It showed a significant difference (Mann–Whitney $U = 116.5$, $Z = -2.0$, $p < 0.05$). Contrary to expectation, the level of distinctness did not differ between the two groups of participants.

Table 1. *Language abilities at pre-test, group means (and standard deviations)*

Measure	Dyslexics	Controls	Difference
Non-word decoding	22.0 (6.1)	31.3 (9.7)	$t = 3.6$, $p < 0.001$
Receptive vocabulary raw score	14.6 (2.3)	13.4 (2.2)	$t = 1.7$, n.s.
Phoneme awareness			
% correct substitutions	58.6 (19.2)	85.0 (7.2)	$t = 5.6$, $p < 0.001$
% intact syllables	78.7 (13.2)	91.0 (3.9)	$t = 3.9$, $p < 0.001$
Quality of phonological repr.s			
Pronunciation accuracy (%)	93.4 (6.4)	96.7 (5.2)	$t = 1.8$, n.s*
Pronunciation distinctness (%)	80.5 (8.1)	82.3 (9.0)	$t < 1$, n.s.

* see text.

Table 2. *Word and non-word learning: paired associate learning of words and non-words with pictures**

Measure	Dyslexics	Controls	Difference
Real words			
Number of trials required	6.7 (3.8)	5.5 (3.4)	$U = 144$, n.s.
Medium-term retention	2.7 (1.1)	2.6 (1.2)	$U = 169$, n.s.
Non-words			
Number of trials required	14.4 (3.0)	10.3 (4.3)	$U = 76$, $p < 0.005$
Medium-term retention	1.6 (0.9)	2.1 (1.4)	$U = 125$, n.s.

* Means (and standard deviations) in each of the two groups with Mann-Whitney *U*-test comparisons of means added in the last column.

Turning towards the measures of verbal learning (Table 2), it was found that the dyslexics were less able to learn new words than the controls. They required more trials (repetitions) to learn to associate the four novel names with pictured persons. Thirteen of 19 dyslexics had not learnt the novel names after 15 trials, whereas only 5 of 19 normal controls failed to learn the novel names after 15 trials. Yet, the dyslexics performed at the level of the controls in the task with real names. There were no differences in the medium-term retention of the taught name-picture associations.

The dyslexics also displayed problems with the acquisition of new *variants* of known words (Fig. 2a–c). The dyslexics progressed significantly less than the controls with the trained words from pre- to post-test ($t(35) = 3.0$, $p < 0.005$, excluding participants with 100% scores at both pre- and post-test), whereas there were no significant group differences in gains with the repeated words or the untrained control words ($t(29) = 1.0$, n.s., and ($t(34) < 1$, n.s., respectively). A 2 groups × 3 conditions (trained, repeated and control words) ANOVA with repeated measures of *gain scores* revealed a significant effect of condition ($F(1, 27) = 3.9$, $p < 0.05$), an insignificant group effect ($F(1, 27) = 3.1$, $p < 0.1$), and an insignificant condition × group interaction ($F(1, 27) = 2.8$, $p < 0.1$). However, planned within-subjects contrasts showed a significant interaction between group and trained words versus control words ($F(1, 27) = 5.8$, $p < 0.05$). This interaction suggests that the group difference in gains with the trained words is not readily explained as a test-retest effect. The pre-test levels

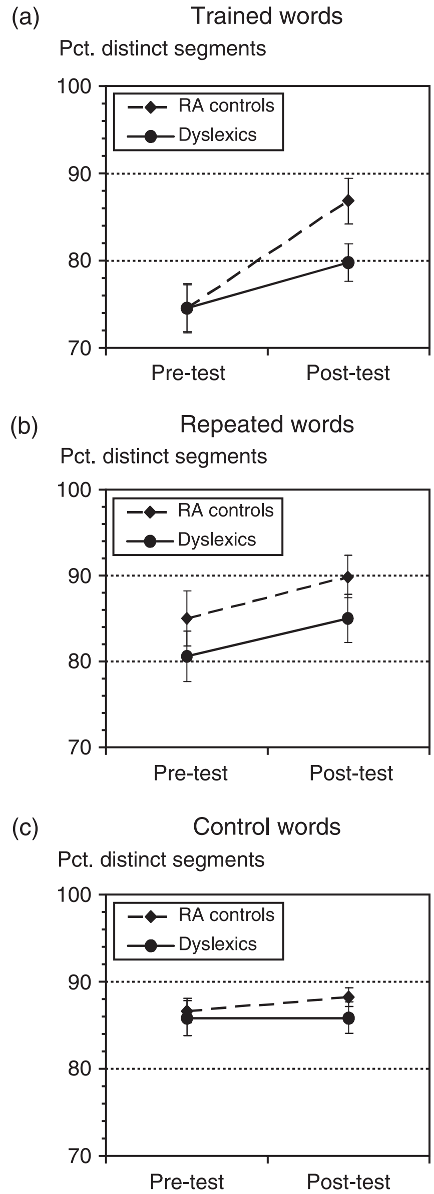

Fig. 2. Learning distinct phonological representations. Percentage of fully distinct vowel segments in three conditions in two groups (error bars represent ± 1 standard error).

Table 3. *Development of phoneme awareness with words in three conditions, pre- and post-test means (and standard deviations)**

Measure	Pre-test	Post-test	Signif. of gain
Dyslexics			
Trained words			
% correct substitutions	54.5 (16.2)	62.8 (20.4)	$t = 3.4$, $p < 0.005$
% intact syllables	78.0 (12.9)	79.0 (11.5)	$t < 1$, n.s.
Repeated words			
% correct substitutions	63.4 (22.0)	65.7 (21.1)	$t = 1.0$, n.s.
% intact syllables	81.5 (15.0)	82.5 (12.0)	$t < 1$, n.s.
Control words			
% correct substitutions	58.0 (23.8)	65.6 (20.7)	$t = 2.5$, $p < 0.05$
% intact syllables	76.8 (16.1)	78.3 (13.5)	$t < 1$, n.s.
Controls			
Trained words			
% correct substitutions	80.1 (7.8)	87.8 (9.4)	$t = 3.2$, $p < 0.005$
% intact syllables	87.8 (4.7)	92.2 (7.1)	$t = 2.8$, $p < 0.05$
Repeated words			
% correct substitutions	88.5 (9.6)	92.1 (6.3)	$t = 1.5$, n.s.
% intact syllables	93.3 (4.7)	94.9 (4.3)	$t = 1.5$, n.s.
Control words			
% correct substitutions	86.5 (9.2)	90.1 (7.7)	$t = 1.6$, n.s.
% intact syllables	91.8 (6.6)	92.5 (7.8)	$t < 1$, n.s.

* Gains with trained words are associated with learning more distinct representations of the same words.

of distinctness did not match completely across the three word conditions. Possible reasons for this are discussed below.

A possible effect of increased quality of phonological representations on phoneme awareness was studied by means of pre- and post-tests of phoneme awareness with distinctness-trained words and control words (Table 3). Looking first at the proportion of correct vowel substitutions, both groups of participants gained significantly in phoneme awareness with the trained words, but apparently less with the untrained control words. While this pattern of results suggest a transfer effect from distinctness to phoneme awareness, the interaction between group and condition was not significant (in a 2 groups × 3 conditions repeated ANOVA on the gain scores from pre- to post-test). So, distinctness training may have resulted in a more general improvement of phoneme awareness across words than expected. There were no significant group differences in gains in any of the three conditions.

For the *other* parts of the syllabic structure, i.e. the consonants, distinctness training appeared to have a positive effect for the RA controls, but not for the dyslexics (Table 3). The RA controls made fewer errors on the consonants (which they were not supposed to substitute) at post-test than at pre-test. The difference between the mean group gains was significant ($t(35) = 2.8$, $p < 0.01$, excluding 2 participants who scored 100% on both occasions). This suggests that training maximally distinct phonological representations was associated with greater gains in consonant stability in the RA control group than in the dyslexic group. A 2 groups × 3 conditions repeated ANOVA on the gain scores from pre- to post-test showed a significant main effect of group ($F(1, 30) = 7.1$, $p < 0.05$), so RA controls gained more than dyslexics across all items. But the analysis showed no significant main effect of condition, and no significant group × condition interaction effect. This suggests that the training of distinct phonological representations was associated with a larger gain in phoneme awareness across words in the dyslexic group than in the RA controls.

DISCUSSION

The first research question was whether dyslexics have less well specified phonological representations than younger normal readers (reading level controls) – even though the older dyslexics may have larger vocabularies. The present study replicated and extended previous results from adults (Elbro *et al.*, 1994; Fowler & Swainson, 1999). The results indicated that the dyslexics had slightly less accurate phonological representations of long, familiar words than younger reading-age controls. The dyslexics pronounced the long, familiar words less accurately, with twice as many errors as the controls did. This finding replicates findings from English (Nation *et al.*, 2001) and extends previous results with the "poor parrot" paradigm into young dyslexics. However, in contrast to previous findings in adults, dyslexic and normal participants pronounced words at the same level of distinctness. One reason may be a ceiling effect on accuracy. Age and experience with written words may also have favoured the dyslexics in the present study.

Some of the results with the phoneme awareness task may shed further light on the phonological representations issue. Because the same words were used to assess both phoneme awareness and quality of phonological representations, it is possible to compare the performance with the two tasks directly. The dyslexics performed significantly worse than the controls on the phoneme awareness task at both pre- and post-test. Both groups performed much worse than would have been expected had their pronunciation accuracy been the only decisive factor. Interestingly, the added processing demands of the vowel substitution task in the phoneme substitution task affected not only the vowels. Even though the task was to substitute only the vowels of the words, the dyslexics produced many responses in which consonants were also erroneously changed. This happened in about 20% of the responses of the dyslexic group – compared with a significantly lower rate of about 9% of the responses in the normal controls (Table 1). This breakdown of the consonantal structure suggests that the phonological representations may be more *fragile* in the dyslexic participants than in the normal controls. Even though all consonantal segments may be present when dyslexics pronounce the words with no distraction, some of the segments may be vulnerable to deletion or change when neighbouring segments are being manipulated. This could be a consequence of less well specified single segments of the phonological representations (cf. the example with *string* in the introduction). However, it should be noted that there was a 2 : 1 proportion between error rates in the dyslexic and the normal group in pronunciation accuracy. This proportion is not far from the error rates in the more demanding phoneme awareness task. A more sensitive measure of pronunciation accuracy would be needed to reliably assess the amount of variance in phoneme awareness that may be accounted for by variance in pronunciation accuracy.

The second research question was whether dyslexics have problems with the acquisition of new verbal (especially phonological) material as compared with younger reading-age controls. The results replicated earlier findings with English, Dutch and German speaking participants (cf. the introduction): the dyslexics took longer to learn to associate novel words (pseudo-names) with pictures of people. Medium term (1 week) retention was not significantly worse, however. This last result was somewhat surprising but replicated previous research (Messbauer & de Jong, 2003).

The third question was a new variant of the second question: do dyslexics' problems with the acquisition of new verbal material extend into poor acquisition of higher quality representations of already known words? The results from the present study suggest that the answer is positive. The dyslexics gained less than normal controls from the distinctness training. They learned less from the imitation task in which they were asked to pronounce known words in the same (distinct) way as a computer-generated voice. This is an important result because it suggests that dyslexia may be associated with problems with the acquisition of phonological material *in general*.

The fourth question was whether word-specific gains in quality of phonological *representations* are associated with gains in phoneme *awareness* with the same words. Distinctness training certainly resulted in significant improvements in phoneme awareness. This was found in both the dyslexic group and the control group. However, an analysis of the stability of the consonant segments in the phoneme substitution task suggested that dyslexics improved less than the RA controls from pre- to post-test. Hence, the gains in phoneme awareness associated with training of the quality of the phonological representations of the words appeared to be larger for the controls than for the dyslexics. Yet, the gains with trained words were not significantly greater than the (non-significant) gains with untrained words. Hence, a non-significant transfer effect to untrained words may be at play which masked the effect of the trained words.

There are limitations of the study which call for some caution in the interpretation of the results. First, the two groups of participants were fairly small, so we may have missed significant differences between groups and training and transfer effects.

Second, the dyslexic adolescents may have had smaller vocabularies than would be expected given their age. It is hard to know for sure because of the lack of a chronological age control group (or a proper standardization of the vocabulary measure). But it is odd that the dyslexics did not score *significantly* higher than the younger controls on the receptive vocabulary measure. So one may speculate that they were not just dyslexic, but more generally language impaired. On the other hand, it should be borne in mind that reading difficulties put severe limitations on vocabulary development (Juel, 1988; Stanovich, 1993). So the dyslexic adolescents may have gradually *acquired* a more widespread delay in language abilities *because* of their dyslexia.

Third, the three sets of words were not completely equated for "difficulty", i.e., pre-test knowledge of the most distinct pronunciation. However, it is hard to match words on this parameter, because there is no simple relationship between well-established parameters, such as frequency, concreteness, imaginableness, etc. and knowledge of the most distinct pronunciation. Frequency of use tends to "wear off" distinctive features of pronunciation. On the other hand, children are more likely to have heard different pronunciation variants of frequent than less frequent words. In any case, the various types of distinctness reductions were fairly evenly distributed across conditions, so it is unlikely that progress with words in each condition would be dependent on word type rather than training condition.

Despite these possible limitations, we think that the study provides some valuable additions to current knowledge about the language bases of reading development and dyslexia. The study adds to the evidence that dyslexia is associated with problems with the acquisition of new phonological material. Not only do dyslexics appear to have problems with the acquisition of phonological representations of new words (pseudo-names), their problems appear to extend to new, more distinct, spoken *variants* of already known words.

Furthermore, increased quality of phonological representations appeared to be less closely linked to improved segmental stability in a phoneme substitution task in the dyslexic adolescents than in the normal controls. All in all, the results are in line with the hypothesis that poorly specified phonological representations may be an underlying problem in dyslexia.

The research reported in this paper was supported by the Danish Research Council for the Humanities (#9700997) and the Nordic Research Councils (contract NOS-S No. 124811/541) to the first author. The authors wish to thank the participating children and adolescents and their teachers from Gerbrandsskolen, Standvejsskolen, Ordblindeskolen in Rødovre, and Ordblindeinstituttet. We are also indebted to Rikke Christensen and Pia Nemholt for help with collection and scoring of data. Two anonymous reviewers deserve thanks for some very constructive comments on an earlier version of the paper.

REFERENCES

Aguiar, L. & Brady, S. (1991). Vocabulary acquisition and reading ability. *Reading and Writing: An Interdisciplinary Journal*, *3*, 413–425.

Allerup, P. & Elbro, C. (1998). Comparing differences in accuracy across conditions or individuals: An argument for the use of log odds. *The Quarterly Journal of Experimental Psychology*, *51A* (2), 409–424.

Borstrøm, I., Petersen, D. K. & Elbro, C. (1999). *Hvordan kommer børn bedst i gang med at læse?* (*How do children best learn to read?*, in Danish). Copenhagen: Centre for Reading Research and The Ministry of Education.

Brady, S. A. (1997). Ability to encode phonological representations: An underlying difficulty in poor readers. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 21–48). Mahwah, NJ: Erlbaum.

Brink, L., Heger, S., Lund, J. & Jørgensen, J. N. (1991). *Den Store Danske Udtaleordbog (The Unabridged Dictionary of Danish Pronunciation)*. Krbenhavn: Munksgaard.

Carney, E. (1994). *A survey of English spelling*. London: Routledge.

Elbro, C., Borstrøm, I. & Petersen, D. K. (1998). Predicting dyslexia from kindergarten. The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, *33* (1), 36–60.

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

Elbro, C. & Pallesen, B. R. (2002). The quality of phonological representations and phonological awareness: A causal link? In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 17–32). Amsterdam and Philadelphia: John Benjamins.

Elbro, C. & Petersen, D. K. (2004). Long-term effects of phoneme awareness and letter name training. An intervention study with children at risk of dyslexia. *Journal of Educational Psychology*. (Revised manuscript submitted).

Elbro, C. & Scarborough, H. S. (2004). Early identification. In P. Bryant & T. Nunes (Eds.), *International Handbook of Children's Reading*. Dordrecht: Kluwer.

Fowler, A. E. & Swainson, B. (1999, April). *Phonological representation of lexical items in good and poor readers: Evidence from three experimental measures.* Paper presented at the biannual meeting of the Society of Research in Child Development, Albuquerque, NM.

Foy, J. G. & Mann, V. (2001). Does strength of phonological representations predict phonological awareness in preschool children? *Applied Psycholinguistics*, *22* (3), 301–325.

Griffiths, Y. M. & Snowling, M. J. (2002). Predictors of exception word and nonword reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology*, *94*(1), 34–43.

Grigorenko, E. L. & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, *124* (1), 75–111.

Grønborg, A., Lund, J., Møller, O. S. & Pedersen, H. (1993). *Ordkendskabstesten* ("The word knowledge test", in Danish). Herning: Special-pædagogisk Forlag.

Johnston, R. S. & Scott, H. (2000, July). *The relationship between phonological working memory and vocabulary acquisition.* Paper presented at the Meeting of the Society for the Scientific Study of Reading, Stockholm, Sweden.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80* (4), 437–447.

Lyon, G. R. (1995). Toward a definition of dyslexia. *Annals of Dyslexia*, *45*, 3–27.

Mayringer, H. & Wimmer, H. (2000). Pseudoname learning by German-speaking children with dyslexia: Evidence for a phonological learning deficit. *Journal of Experimental Child Psychology*, *75*, 116–133.

Messbauer, V. C. S. & de Jong, P. F. (2003). Word, nonword and visual paired associate learning in Dutch dyslexic children. *Journal of Experimental Child Psychology*, *84* (2), 77–96.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91* (1), 3–19.

Nation, K., Marshall, C. M. & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skills: Evidence from children with developmental reading disorders. *Language and Cognitive Processes*, *16* (2/3), 241–259.

Pennington, B. F., Lefly, D. L., Van Orden, G. C., Bookman, M. O. & Smith, S. D. (1987). Is phonology bypassed in normal or dyslexic development? *Annals of Dyslexia*, *35*, 62–89.

Petersen, D. K. (2000). What do you get if you add mmm/ to ice? Training phoneme awareness in kindergarten. An intervention study with children of dyslexic parents. In N. A. Badian (Ed.), *Preschool prediction and prevention of reading failure* (pp. 247–273). Baltimore, MD: York Press.

Petersen, D. K. & Elbro, C. (1999). Pre-school prediction and prevention of dyslexia: A longitudinal study with children of dyslexic parents. In T. Nunes (Ed.), *Learning to read: An integrated view from research and practice* (pp. 133–154). Dordrecht: Kluwer.

Rack, J., Hulme, C., Snowling, M. & Wightman, J. (1994). The role of phonology in young children learning to read words: The direct-mapping hypothesis. *Journal of Experimental Child Psychology*, *57*, 42–71.

Snowling, M., Goulandris, N., Bowlby, M. & Howell, P. (1986). Segmentation and speech perception in relation to reading skill: A developmental analysis. *Journal of Experimental Child Psychology*, *41*, 489–507.

Stanovich, K. E. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. In H. Reese (Ed.), *Advances in child development and behavior* (Vol. 24, pp. 133–180). Orlando, FL: Academic Press.

Stone, B. & Brady, S. (1995). Evidence for deficits in basic phonological processes in less-skilled readers. *Annals of Dyslexia*, *45*, 51–78.

Swan, D. & Goswami, U. (1997). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of Experimental Child Psychology*, *66* (1), 18–41.

Vellutino, F. R., Steger, J. A., Harding, C. J. & Phillips, F. (1975).

Verbal vs. non-verbal paired-associates learning in poor and normal readers. *Neuropsychologia*, *13*, 75–82.

Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, *13*, 286–350.

Wesseling, R. (1999). *Quality of phonological representations: Etiology and remediation of dyslexia*. Amsterdam: Free University. Doctoral dissertation.

Wesseling, R. & Reitsma, P. (2001). Preschool phonological representations and development of reading skills. *Annals of Dyslexia*, *51*, 203–229.

Wimmer, H., Mayringer, H. & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology*, *92* (4), 668–680.

Received 18 August 2003, accepted 2 April 2004

APPENDIX

The words listed below were the material in the static and dynamic tests of quality of phonological representations and phoneme awareness. The boldfaced vowel segments are those which are the most likely to be reduced in standard pronunciation and the ones used in the scoring of the distinctness level of the phonological representation.

Trained	*Repeated*	*Control*
chok**o**lad**e** ("chocolate")	krok**o**dill**e** ("crocodile")	frik**a**dell**e** ("rissole")
hel**i**kopter ("helicopter")	pol**i**timand ("policeman")	lok**o**motiv ("locomotive")
lykk**e**lig ("happy")	hygg**e**lig ("cosy")	virk**e**lig ("really")
flyv**e**maskin**e** ("aeroplane")	flød**e**boll**e** ("cream puff")	børn**e**hav**e** ("kindergarten")
kirs**e**bFr ("cherry")	mærk**e**lig ("strange")	kirk**e**tårn ("church tower")
skol**e**task**e** ("school bag")	brill**e**ab**e** ("four-eyes")	bad**e**hætt**e** ("bathing cap")
bibl**i**otek ("library")	mat**e**m**a**tik ("math")	pap**e**gøj**e** ("parrot")
tygg**e**gummi ("chewing gum")	slikk**e**pind ("lollipop")	sjipp**e**tov ("skipping rope")
skrald**e**spand ("waste bin")	trill**e**bør ("wheel barrow")	tryll**e**stav ("magic wand")