British Journal of Educational Psychology (1999), **69**, 587–605 *Printed in Great Britain*

Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners

Morag Stuart*

Psychology and Special Needs Group, Institute of Education, London University

Background. Previous studies demonstrate that phoneme awareness training, particularly when combined with letter-sound teaching, results in improved reading and spelling development.

Aims. This study seeks to extend previous findings by (a) including children learning English as a second language, who have typically been excluded from previous studies; (b) providing training for whole classes, rather than small groups; (c) using a commercially available programme; and (d) giving minimal training to teachers administering the programme.

Sample. Two groups ($N = 112$) of 5-year-olds, 96 of whom were learning English as a second language, were enrolled into either the experimental (phoneme awareness and phonics) programme or the control programme, which took a more holistic approach based on Holdaway's (1979) use of Big Books.

Method. Children were pretested on measures of spoken and written language, phonological awareness and alphabet knowledge, prior to a 12-week intervention using either the experimental or control programme. Children were post-tested on all measures immediately after intervention, and again one year later.

Results. The experimental programme accelerated children's acquisition of phoneme awareness and of phonics knowledge, and their ability to apply these in reading and writing. In the year following intervention both groups made comparable progress in most areas; however, at the end of this year the experimental group were still significantly ahead in phoneme awareness and phonics knowledge, and on standardised and experimental tests of reading and spelling.

Conclusions. Early concentration on teaching phoneme awareness and phonics can radically improve reading and spelling standards in inner city second language learners.

*Requests for reprints should be addressed to Dr Morag Stuart, Psychology and Special Needs, Institute of Education, London University, 25 Woburn Square, London WC1H 0AA, UK.

The National Literacy Strategy (1998) requires schools and teachers in England and Wales to incorporate phonological awareness training into the early reading curriculum. Three aspects of this training deserve comment.

First, the phonological awareness training set out by the NLS includes both onset-rime and phoneme awareness. This catholic approach reflects the continuing debate amongst researchers concerning the nature of the phonological units first available to and used by children in reading – are these onset-rime units (e.g., Bryant, 1998; Goswami & Bryant, 1990; Maclean, Bryant & Bradley, 1987) or phonemes (e.g., Duncan, Seymour & Hill, 1997, in press; Hoien, Lundberg, Stanovich & Bjaalid, 1995; Hulme, Muter & Snowling, 1998; Muter, Hulme, Snowling & Taylor, 1997)?

Second, phonological awareness training is to be accompanied by teaching how to represent speech sounds with letters. This emphasis on immediately linking sound patterns with their corresponding spelling patterns is soundly based in recent research findings. Thus, although numerous training studies in the literature substantiate the claim that early phonological awareness training has beneficial effects on reading development (e.g., Cunningham, 1990; Lundberg, Frost & Petersen, 1988), most training studies include some explicit instruction in the connections between sound segments and letters of the alphabet, with some evidence that this enhances effectiveness (e.g. Ball & Blachman, 1991; Bradley & Bryant, 1983; Hatcher, Hulme & Ellis, 1994), and that both the phonological awareness and the letter-sound instruction are necessary components of training (e.g. Ball & Blachman, 1991; Byrne & Fielding-Barnsley, 1991).

Third, phonological awareness and letter-sound training in the NLS lead in to a structured phonics teaching programme. The training studies carried out to date have little to say concerning the necessity for following such early training with two years of structured phonics teaching as proposed by the NLS. However, naturalistic studies of children who arrive in school as 5-year-olds with good phonological awareness and letter-sound correspondence knowledge suggest that these assets may be sufficient to drive development of a self-teaching printed word recognition system (Share, 1995). For example, preschool 4-year-olds with good phonological awareness continued to show significantly better reading of regular and exception words and nonwords as 10-year-olds than their peers who had started school unaware that spoken words were patterns of sound (Stuart & Masterson, 1992). Stuart, Masterson, Dixon & Quinlan, (1999) found that 6- and 7-year-olds taught only single letter-sound correspondences at school had inferred both vowel digraph correspondences and rime spellings from their reading experience. Stuart *et al.* suggest that, as phonologically aware children develop proficiency in reading, their sight vocabulary becomes a data base from which context-specific spelling patterns are inferred. Similar suggestions have been put forward by Share (1995), Thompson, Cottrell & Fletcher-Finn (1996) and Zinna, Liberman & Shankweiler (1986). Since there is at present no clear evidence that these findings will generalise to children given the kinds of early explicit instruction outlined in the NLS, it is probably wise at this point to recommend continuing teaching – but the necessity for this should be evaluated as the programme proceeds. It is intended that follow-up investigations of the children in the present study will shed light on this issue.

The present study set out to replicate and extend previous research findings of the beneficial effects of phonological awareness and letter-sound correspondence training.

Four particular extensions were made. First, although two previous studies have already demonstrated the effectiveness of such training for inner city children from socially deprived backgrounds (Blachman, Ball, Black & Tangel, 1994; Brady, Fowler, Stone & Winbury, 1994), neither of these included a control condition, and children learning English as a second language were specifically excluded. In the present study, proper controls were implemented, and 86% of the children were English second language learners; many of them entered school with limited ability to express themselves in English. However, recent research suggests that children exposed to more than one phonological system are likely to have heightened levels of phonological awareness (Bruck & Genesee, 1995; Campbell & Sais, 1995). Thus, introducing reading through a programme of phonological awareness and letter-sound teaching should capitalise on the strengths of second language learners, whereas the now traditional holistic approaches emphasising meaning would seem to focus on their potentially weakest area.

Second, most previous studies have either involved research staff in delivering the training programmes (e.g., Byrne & Fielding-Barnsley, 1991), or have required extensive pre-training of school staff involved in delivery (e.g., Brady *et al*, 1994; McGuiness, McGuiness & Donohue, 1995). In the present study, the children's classroom teachers delivered the programme having received minimal pre-training, which perhaps more accurately reflects conditions in the real world, but which, in a comparative study by Byrne & Fielding-Barnsley (1995), resulted in rather less promising results than their earlier researcher-delivered training.

Third, in most previous training studies children have been taught in small groups containing four to six children. In the present study, the whole class of children was taught together, as recommended in the National Literacy Strategy Framework for Teaching (p. 11). However, contrary to the recommendations of the NLS, only phoneme awareness was trained, with no attempt to introduce rhyme or onset and rime units.

Fourth, in most previous studies, researchers have developed and implemented teaching programmes they themselves have designed. The present study used an off-the-peg commercially available teaching programme designed by a teacher (Lloyd, 1992), which has already been shown to have positive effects on developing reading and writing skills (Johnston & Watson, 1997; Watson & Johnston, 1999). This affords the possibility for interested practitioners easily to gain access to the materials and procedures employed and try them out in their own classrooms.

The following questions were investigated: (1) is it possible to accelerate phoneme segmentation skills and knowledge of grapheme-phoneme correspondences, in inner city children many of whose first language is not English, by using a whole class teaching procedure implemented by classroom teachers over a 12-week period in the first year of formal schooling? (2) If so, does this acceleration lead to more successful development of reading and spelling skills at the end of the second year at school?

Participants

A total of 122 children (mean age at pretest 5 years 0 months, SD 3 months) were enrolled into this longitudinal study: data are reported on the 112 children present throughout the study. When pretesting started in January 1997, most of the children were attending

Early Years Units containing both 4- and 5-year-olds. Only 16 of the 112 children spoke English as their first language; the vast majority of the second language learners were Sylheti speakers, with three Cantonese speakers, and four speakers of other languages.

Interventions

Primary schools within the (sometimes tightly defined) areas of benefice of the several charities and institutions funding the project were approached and asked if they would be willing to take part in a study comparing two different approaches to the early teaching of reading. Each approach was briefly outlined; schools expressing interest in taking part were asked which intervention they preferred to implement, in order to maximise teacher motivation to carry out the intervention to the best of their ability, in an attempt to demonstrate that their chosen method was the best. One class in each of three schools undertook to implement the 'Big Books' (henceforward BB) intervention. Three classes from two further schools undertook to implement the 'Jolly Phonics' (henceforward JP) intervention. Since schools were allowed to opt for one or other intervention, it is important to show that schools were, nevertheless, similar across interventions. As shown in Table 1, when statistics detailing the social, ethnic and linguistic composition of each school and their recent Key Stage 1 SAT's performance were compared across schools in each intervention, no significant differences were found. Data are reported on 57 children (mean age at pretest, 5 years 1 month) in the BB intervention, and 55 children (mean age at pretest, 5 years 0 months) in the JP intervention.

Training advice to teachers

Teachers undertaking the BB intervention were accustomed to using these large (A1 size) books with their classes. The researcher met each teacher, and discussed the different kinds of use which could be made of big books. Teachers were particularly asked to spend time on word level work, that is, to emphasise words and letters, by drawing children's attention to written words in the text, and talking about the letters in words. Work with letters should involve introduction to their names and sounds, and children should be encouraged to notice and learn words and letters in the classroom environment. Activities to foster word and letter learning were discussed: the teachers were already using many imaginative and fun activities to these ends, such as having a collection of handbags, each with a different letter, containing small interesting objects whose names began with the sound of the letter.

Teachers undertaking the JP intervention were asked to make a rather more radical adjustment to their usual teaching methods. The researcher met each teacher and discussed the contents of the programme. Teachers were given the Phonics Handbook to read, which describes the programme, its rationale, and the materials provided to support teaching. They were also given copies of a training video with footage of teachers using the materials, and offered the opportunity to attend a training seminar given by one of the authors of the programme: two of the three teachers attended.

Materials supplied to schools
Teachers in the BB intervention were invited to submit a list of new books up to a ceiling of £200, which were then ordered and bought for them. The big books provided in this way were accompanied, where available, with sets of smaller copies for children to read together. One teacher had planned to focus on particular authors during the intervention period, and attractive big book copies of these authors' work were made for her to use.

Teachers in the JP intervention were supplied with one box of the complete programme materials: the Phonics Handbook, containing worksheets which were photocopied for them; one set of phonics workbooks; one set of finger phonics books; one phonics wall frieze; one set of jiglets (six simple jigsaw puzzles with a picture and a word on each); one set of stencilets (stencils which children can use to draw a picture and write the word for it); and two children's videos, in which Inky Mouse shows Snake and Bee how to read and write. Total cost of each pack was £130, with some additional photocopying costs.

Intervention
The intervention period was 12 weeks. Teachers were asked to spend one hour each day on reading and writing, either centred around big books, or using the phonics programme. Each hour began with whole class activities; children then split into groups for follow-up activities. In most cases, the hour was an unbroken stretch of time first

Table 1. Characteristics of schools in each intervention

Measure	mean % (SD)	
	JP	BB
3 terms in nursery	29.4 (19.8)	22.4 (15.5)
6 terms in nursery	52.0 (17.5)	53.9 (9.8)
KS1 level 1	39.6 (2.3)	39.1 (9.4)
KS1 level 2	45.4 (9.3)	46.4 (11.7)
KS1 level 3	15.0 (7.1)	14.8 (16.9)
Fluent ESL	2.6 (3.7)	7.3 (7.8)
Stage 3 ESL	7.7 (9.3)	13.5 (1.8)
Stage 2 ESL	18.7 (7.2)	29.2 (5.2)
Beginner ESL	52.9 (40.5)	29.8 (11.1)
Free meals	70.8 (9.6)	69.0 (10.9)
Attendance	92.7 (1.4)	88.6 (1.8)
Large families	49.3 (7.9)	47.9 (8.8)
One parent families	9.3 (8.8)	14.5 (6.4)
	mean (SD)	
Pupil mobility	9.6 (3.5)	8.2 (3.5)
EPI*	10.3 (1.9)	9.2 (1.4)
Pupil teacher ratio	18.4 (1.5)	19.2 (3.9)

No differences reached statistical significance at $p < .05$
* EPI = Educational Priority Index

thing in the morning; however, one teacher preferred to split the hour, with half an hour in the morning devoted to whole class activity, and another half hour in the afternoon for group and individual follow-up activities. Teachers were visited regularly by the project research assistant, who was able to confirm that all children were receiving each intervention daily for the prescribed amount of time. At the end of the intervention period, teachers in both interventions expressed the view that their children had profited from the changes and adaptations introduced.

Timetable

Prior to intervention, children were pretested in January and February, 1997 on nine experimental measures of phoneme awareness, phonic knowledge, reading and writing, and nine control measures of oral language, auditory perception, rhyme awareness, alphabet knowledge, and mathematics. Intervention took place from March to May. Children were post-tested on all measures in June and July, 1997, immediately following intervention. Delayed post-tests took place one year later, in June and July 1998. All pre- and immediate post-test measures were repeated, with some additional tests of reading, spelling and writing to dictation. On all testing occasions, each child was seen four times in sessions that lasted up to 25 minutes. Tests were given to all children in the same predetermined order, which was designed to provide variety and ensure that in each session children were given a mix of easy and harder tasks.

Control measures

(a) Oral language

As oral language abilities obviously influence literacy outcomes, it was important that children should be matched across interventions for oral language skills. We therefore measured the children's receptive vocabulary using the 2nd edition of the British Picture Vocabulary Scales (BPVS), (Dunn, Dunn, Whetton & Burley, 1997) which gives norms for children learning English as an additional language, as well as for English children, and was therefore particularly suitable for the present study. We also took two measures of the children's syntactic abilities. We used the Linguistic Concepts subtest of the Clinical Evaluation of Language Fundamentals – Preschool (CELF) (Wiig, Secord & Semel, 1992), which evaluates children's ability to comprehend oral directions that (a) contain early acquired linguistic concepts such as co-ordination, inclusion/exclusion, temporal relations; (b) involve quantifiers and ordinals; and (c) increase in length from one- (e.g., 'Point to one of the bears') to three-level commands (e.g., 'Point to the bear, the turtle and the fish'). Children are taught the names for animals used in the test booklet prior to administration of the 20 test items. We also used a Sentence Repetition task (Willows, 1996) in which children were asked to repeat 10 sentences varying in length and syntactic complexity. Where children do not have a syntactic structure available, they will simplify the structure of a complex sentence in their repetition.

(b) Auditory perception

As we were going to train and compare performance on phonological tasks, we needed to be sure that the children in each intervention had equally good auditory perception. We therefore gave them a Phoneme Discrimination test (Willows, 1996) containing two training and 14 test items. Children were asked to indicate whether pairs of sounds spoken (with mouth hidden) by the tester were the same (e.g., /k/../k/) or different (e.g., /k/../t/). Children failing to respond correctly to the first two test items were shown pictures that were either the same (e.g., two butterflies) or different (e.g., a butterfly and a car) to check that they understood the task requirements and/or to train them to understand 'same' and 'different'.

(c) Alphabet knowledge
As we were going to train and compare children's knowledge of letter-sounds, we included letter name knowledge as a control measure for non-specific improvements in letter knowledge. Letter name recognition was tested by asking children to point to the 26 letters of the alphabet, presented in a booklet with 9–9–8 letters per page, as the tester named them in non-alphabetic order. Letters commonly confused by young children (e.g., d/b, w/m, u/n) appeared on different pages. Letter name recall was tested by the tester pointing in turn to each letter (with letters presented in non-alphabetic order in a booklet with 9–9–8 letters per page) and asking the child to say its name.

(d) Rhyme awareness
As we were going to train and compare children's phonemic awareness, we included a rhyme awareness measure to control for non-specific improvements in phonological awareness. We used a Rhyme Detection test (Stuart, 1995) with 4 practice and 12 test items. Children were shown sets of three pictures of familiar objects whose names either rhymed (e.g., 'jar, car, star') or did not rhyme (e.g., 'gate, horse, leaf'). Children were asked to indicate whether the names rhymed or not.

(e) Mathematics knowledge
We included the British Ability Scales (BAS) Basic Number Skills subtest (Elliott, Murray & Pearson, 1983) as a check on the specificity of intervention effects. If these are specific to literacy skills, then at post-tests children should differ in literacy skills but not in mathematics knowledge. If at post-tests children differ also in mathematics knowledge, then some non-specific 'Hawthorne effects' may be operative.

Experimental measures
(a) Phoneme awareness
Two tests of phoneme awareness were given. In the Initial Phoneme Identification test (Stuart, 1995) children were shown a set of 24 pictures of objects whose names began with 24 different phonemes, corresponding to the 24 letters of the alphabet that remain when x and q are omitted (e.g., /ɛ/ for 'engine'; /f/ for 'foot'). Following presentation of three practice items, children were asked to name each picture and say what sound the name began with. We also gave an adaptation of the Phoneme Segmentation test (Yopp, 1988), presenting only 11 of the 22 test items to avoid unduly stressing the children. All five practice items were used to show the children how to break words apart into their sounds (e.g., to repeat 'cat' as '/k/../ae/../t/').

(b) Phonic knowledge
This was tested in three ways. Letter sound recognition was tested by asking children to point to the letters that represent each of the 26 phonemes (one for each letter of the alphabet) spoken by the tester. Letters were presented in non-alphabetic order in a booklet with 9–9–8 letters per page, and letters which can represent the same sound

(e.g., 'k' and 'c') appeared on different pages. Letter sound recall was tested by showing children a card with 43 graphemes (single letters and digraphs) that represent 43 phonemes in English. The tester pointed to each grapheme in turn and asked children to say its sound. Writing sounds to dictation (Willows, 1996) was tested by asking children to write the letter or letters representing 10 phonemes spoken by the tester. There were six single-letter responses, and four digraph responses.

(c) Reading measures

Reading was tested using two standardised and two experimental tests. In the British Ability Scales (BAS) Single Word Reading Test (Elliott *et al.*, 1983) children were shown a card with words on and asked to read as many words as they could. To avoid undue stress, at pretest the first 10 words only were presented routinely to all children. Children who successfully read some of these words were shown the test card itself. At both post-tests, the word card was presented to all children. In the Young's Group Reading Test (Young, 1985) children were first shown a sheet of pictures, each accompanied by from three to five written words, and asked to circle the word which correctly represented the picture. Children who successfully identified most of the correct words were then asked to choose the missing word (from an array of six words) for incomplete written sentences. Children were tested individually. In the Read Words test to measure early word reading, children were asked to read aloud 14 high frequency words. Seven were two- to four-letter regular words (e.g., 'at', 'big', 'jump') and seven were one- to four-letter irregular words (e.g., 'I', 'to', 'you', 'have'). If the child could not read the word correctly, the tester asked 'Can you sound out the word?'. Responses were scored as incorrect, correct immediately, or correctly analysed. In the Read Nonwords test to measure early phonological recoding ability, children were asked to read 10 CVC nonwords presented in a booklet illustrated with friendly monsters, one per page. Each monster had a CVC name (e.g., 'vus', 'naz') which children (following three practice items) were asked to read: 'I wonder if you can tell me what this monster is? It's a . . . ?'.

(d) Writing measures

We used the Write Words to Dictation test (Willows, 1996) in which children were asked, 'Now, can you write some words for me? Can you write the word (e.g., 'and') for me?' Ten two-, three- and four-letter words were dictated.

(e) Delayed post-tests

At delayed post-test, three additional experimental measures of reading, writing and spelling were included. In the Neale Analysis of Reading Ability – Revised (Neale, 1997) children were asked to read a series of illustrated short prose passages of increasing difficulty and answer questions about each passage. Testing stopped when children reached a predetermined number of errors in reading a passage. Separate measures of reading accuracy and reading comprehension were obtained. As the test is designed for children with reading ages of 6 or above, it could not be used at pretest or immediate post-test. In the Schonell Spelling test (Schonell, 1985) a set of 100 words of increasing difficulty was dictated to children, each first as a single word, then repeated in disambiguating sentence context and finally singly. Testing was stopped after 10 consecutive errors. Again, it was considered unethical to ask children to attempt this difficult task at pretest and immediate post-test, although some of the words which *were* dictated to children on these occasions occur as early items in the Schonell test. In the Clay Dictation test (Clay, 1979) two short sentences were dictated for children to write. Children were encouraged to listen for the sounds they heard in the words. Each sound correctly transcribed scored one point. As this test was designed for use with 6-year-olds it was not considered suitable at pretest and immediate post-test.

Results

Ideally, children in each intervention group should have been well-matched at pretest on all experimental and control measures, and well matched throughout the study on all control measures. However, when pretest results were analysed, it became apparent that a perfect pretest match had not been achieved on all measures, and that, on some of the critical experimental measures (Initial Phoneme Identification; Letter Sound Recognition; Letter Sound Recall, and Writing Sounds to Dictation), the JP group were significantly ahead of the BB group. It was therefore necessary in analysing post-test results to control statistically for pretest differences. One commonly used method for doing this is to covary out pretest scores before analysing the post-test scores, but this method depends on use of analysis of variance, which requires normally distributed data sets. Many of the distributions on measures in the present study departed wildly from normal at both pretest and immediate post-test, with floor effects apparent on both occasions. Thus, many results were simply unanalysable in this way. It was therefore decided to analyse gain scores from pretest to immediate post-test, and from immediate to delayed post-test. The logic behind this decision was that the hypothesis of interest is that the JP intervention should accelerate phoneme awareness and phonics knowledge, that comparison of gain scores should capture any such acceleration, and that acceleration should happen regardless of starting point. Use of gain scores should also jeopardise rather than enhance support for the prediction of accelerated progress in those tests where JP children were already significantly ahead at pretest, since these have finite sets of items. The JP children therefore have less opportunity to improve their scores, as they are already closer to the ceiling.

Where scores were normally distributed and variances homogeneous, results were analysed in one-way ANOVA. Where normality and homogeneity were not present, results were analysed using Mann-Whitney tests. Mean differences over time and their statistical significance are shown in Tables 2–5. In each table, pretest results, when the children's mean chronological age was 5 years, are shown under columns labelled T1 (time 1). Gain scores from pretest to immediate post-test (T2, mean chronological age 5 years 6 months) are shown under columns labelled T1–T2. Gain scores from immediate to delayed post-test (T3, mean chronological age 6 years 6 months) are shown under columns labelled T2–T3. Absolute scores at T3, when the children were at the end of their second year in school, are shown under columns labelled T3.

For control measures, there should be no differences between groups on any measure at any time. For experimental measures, there should be no differences between groups on any measures at pretest. If the intervention has been effective, then the JP group should make significantly more progress from T1–T2 on all measures than the BB group. There should be no differences between groups on any measures from T2–T3, when all schools were continuing to teach their normal programmes. If the early intervention has made a lasting difference to performance, then at T3, the JP group should be significantly better than the BB group on all measures.

Results are presented first for control and then for experimental measures.

Control measures

(a) Oral language

As shown in Table 2, children were well matched at pretest (T1) on the two syntactic measures (Linguistic Concepts, Sentence Repetition). Children in both interventions made equal gains on both these measures during the intervention (T1–T2) and in the following year (T2–T3). At the end of their second year in school (T3), they were still well matched on both measures.

Table 2. Control measures 1: Mean differences in oral language over time

Measure	*N* items		(pretest) T1		Gains T1–T2		Gains T2–T3		(del. post-test) T3	
			JP	BB	JP	BB	JP	BB	JP	BB
Linguistic Concepts	20	mean	9.19	9.73	3.15	3.43	3.2	2.7	15.29	15.86
		SD	(5.4)	(5.0)	(3.9)	(3.7)	(3.2)	(2.7)	(3.3)	(3.4)
			$F = 0.28$ n.s.		$F = 0.14$ n.s.		$F = 0.70$ n.s.		$F = 0.80$ n.s.	
Sentence Repetition	10	mean	2.91	3.00	0.65	0.56	2.2	2.1	5.73	5.57
		SD	(2.7)	(2.7)	(1.9)	(1.8)	(2.5)	(2.2)	(2.7)	(3.1)
			$F = 0.03$ n.s.		$F = 0.08$ n.s.		$F = 0.11$ n.s.		$F = 0.08$ n.s.	
BPVS (English ss)		mean	76.36	81.54	5.46	4.30	5.5	1.9	85.85	87.96
		SD	(14.9)	(12.8)	(11.3)	(10.5)	(10.9)	(10.4)	(11.4)	(13.5)
			$F = 3.69$ n.s. ($p < .06$)		$F = 0.29$ n.s.		$F = 3.18$ n.s. ($p < .08$)		$F = 0.79$ n.s.	
BPVS (ESL ss)		mean	83.90	91.71	6.69	4.48	6.6	0.4	96.13	97.04
		SD	(14.7)	(11.8)	(12.6)	(11.9)	(12.9)	(11.6)	(13.9)	(12.2)
			$F = 7.89$ $p < .01$		$F = 0.71$ n.s.		$F = 5.82$ $p < .02$		$F = 0.12$ n.s.	

When oral vocabulary was assessed against the English test norms, there were again no significant differences at pretest, in gains during or following intervention, or at the end of the second year in school. However, the pretest mean difference approached significance, with children in the BB group having the advantage. In the year following intervention, the mean difference in gains again approached significance, with children in the JP group showing signs of catching up.

When oral vocabulary was assessed against the ESL norms (and with the 16 English first language children dropped from the analysis), the pretest difference was significant. However, as this difference was in favour of the BB group, any influence from superior oral language ability should favour this control group, rather than the experimental JP group. Similarly, the mean difference in gains in the year following intervention was also statistically significant, with the JP children catching up on the BB children's earlier advantage. Thus, by the end of the second year in school, there was again no significant difference between the two groups.

These differences in oral vocabulary development, which are significant when only ESL children are considered against the ESL norms, are not easily interpreted at this point, and will be considered in the discussion when all results have been presented. We can however be confident that any influence from oral vocabulary to reading development should favour the control rather than the experimental group.

It is notable that, despite having experienced from three to six terms of nursery education, the children's mean score on the BPVS was 1–2 standard deviations below the English norms and, when the 16 English first language children were excluded and ESL norms applied, mean scores were still almost one standard deviation below the population mean. This may be an inevitable consequence of the fact that the vast majority of the children shared a common first language, and that this was therefore probably the main language they used in nursery and school, as well as at home.

(b) Auditory perception

As shown in Table 3, the two groups were well matched on the phoneme discrimination test used to measure auditory perception at pretest. They made similar gains during and following intervention, and were still well-matched (and at ceiling) on this test at the end of their second year in school. Good control was therefore achieved throughout the study.

(c) Alphabet knowledge

Table 3 also shows that children were well matched for alphabet knowledge at pretest, whether this was tested in a recognition or a (harder) recall task. When letter name knowledge was tested in a recognition task there was no difference in gains made during or following intervention, or at the end of the second year in school. However, when this knowledge was tested in a recall task, the JP group made significantly more improvement during the intervention, but not in the following year, and the groups were well matched again at the end of the second year in school. This suggests that the JP intervention possibly served to improve some aspects of letter name knowledge: when children were required to retrieve letter names from memory, the JP children did better immediately after the intervention. At all other points, alphabet knowledge was well controlled across groups.

(d) Rhyme awareness

On the forced choice rhyme detection test used to assess rhyme awareness, children could score 50% correct by chance. These scores were therefore treated as categorical data (significantly above chance; at chance) and analysed using chi square tests. As shown in Table 3, there were no significant differences between groups at pretest in the percentages of children scoring above chance. There were also no significant differences in the percentages of children who moved from at chance to above chance either during or in the year following intervention, and no significant differences in percentages of children above chance at the end of the second year in school. Thus, good control was achieved over rhyme awareness throughout the study. In both groups, percentages of children scoring above chance improved steadily, so that almost 80% were above chance by the end of the second year in school.

(e) Maths knowledge
As shown also in Table 3, children were well matched for maths knowledge at pretest. During the intervention, children in the BB group made significantly more progress in maths than the JP group. The fact that the greater improvement was found in the BB group does leave intact the specificity of predicted differential intervention effects: they certainly did not transfer to maths in the JP group. In the year following intervention, progress was equal across both groups, and there were also no significant differences between groups at the end of the second year in school.

Summary of results from control measures
Overall, good control was achieved on oral language at pretest, with some differential rates of progress between the ESL children in the two groups in oral vocabulary development. Good control was achieved on auditory perception and on the untreated phonological measure of rhyme awareness throughout the study. Good control was achieved on alphabet knowledge at pretest, and, when this was tested in an easier recognition task, similarity between groups in terms of progress and achievement was maintained throughout the study. When alphabet knowledge was tested in a recall task, there seems to have been an influence from the JP intervention. This was not predicted, as the teachers were not specifically instructed to teach letter names. It seems plausible

Table 3. Control measures 2: Mean differences in alphabet (letter name) knowledge, auditory perception, rhyme awareness and maths knowledge over time

Test	*N* items		(pretest) T1		Gains T1–T2		Gains T2–T3		(del. posttest) T3	
			JP	BB	JP	BB	JP	BB	JP	BB
Phoneme discrim'n	14	mean	11.41	10.74	1.31	1.36	0.6	1.0	13.29	13.02
		SD	(2.9)	(2.9)	(1.3)	(2.9)	(2.2)	(2.1)	(2.1)	(1.7)
			$z = 1.52$ n.s.		$F = .008$ n.s.		$F = 1.19$ n.s.		$F = 0.56$ n.s.	
Letter name recog'n.	26	mean	15.07	14.21	7.25	5.89	2.9	4.8	25.15	25.14
		SD	(9.3)	(9.4)	(7.2)	(6.4)	(5.5)	(5.9)	(2.9)	(2.3)
			$F = 0.24$ n.s.		$F = 0.76$ n.s.		$F = 3.1$ n.s.		$F = 0.0001$ n.s.	
Letter name recall	26	mean	11.92	12.47	9.42	5.98	3.3	5.3	24.71	24.07
		SD	(9.2)	(8.8)	(7.4)	(5.3)	(5.8)	(6.2)	(3.7)	(3.3)
			$F = 0.1$ n.s.		$F = 7.16$ $p < .01$		$F = 2.92$ n.s.		$F = 0.91$ n.s.	
Rhyme detection	above chance		18.5%	32.1%	+ 33%	+ 26.7%	+ 27.2%	+ 19.2%	78.2%	77.2%
			$\chi^2 = 2.7$ n.s.						$\chi^2 = 0.01$ n.s.	
BAS number (raw)		mean	12.43	11.16	4.15	6.11	7.39	6.13	23.67	22.89
		SD	(5.8)	(6.5)	(3.9)	(4.0)	(4.0)	(4.3)	(5.8)	(2.7)
			$F = 1.16$ n.s.		$F = 6.55$ $p < .02$		$F = 2.51$ n.s.		$F = 0.43$ n.s.	

to suggest that the concentrated focus on letter-sound teaching inevitably resulted in more exposure also to letter names. Good control was also achieved overall on maths knowledge, with the only significant difference found here favouring the BB group, and therefore unproblematic in terms of using this measure to assess specificity of intervention effects, since it was the JP intervention that was predicted to enhance literacy development.

Experimental measures

(a) Phoneme awareness

As stated earlier, on the Initial Phoneme Identification test, JP children performed significantly better at pretest than BB children, necessitating the use of gain scores in subsequent analyses to take account of these pretest differences. As shown in Table 4, although the mean difference in gains during the intervention was in favour of the JP group, this difference was not statistically significant. In the year following the intervention, the BB group made significantly more progress than the JP group. This is perhaps inevitable given that the JP group's mean score at T2 was already approaching ceiling (20.5/24): they had little further room for improvement from T2 to T3. However, at T3, the JP group were significantly ahead of the BB group, and the BB group's T3 scores were statistically indistinguishable from the JP group's T2 scores (one year earlier) on this test (mean BB T3 = 19.57, SD 6.7; mean JP T2 = 20.27, SD 6.5; F = 0.31, n.s.). Thus, the BB group never caught up on the early lead of the JP group, even though this early lead cannot be attributed solely to the intervention.

Table 4 also shows that on the more difficult Phoneme Segmentation test, groups were well matched, and at floor, at pretest, and differences over time followed the pattern predicted. The JP group made significantly greater gains during the intervention. In the year following intervention, both groups made equal progress. This is important, as equal progress when each school was following its normal teaching programme suggests that teaching as well as schools and children was well matched. Nevertheless, the effects of early intervention are still apparent at the end of the second year in school, with JP children significantly ahead of BB children at T3 and the BB group's T3 scores again statistically indistinguishable from the JP group's T2 scores (one year earlier) on this test (mean BB T3 = 4.67, SD 3.7; mean JP T2 = 4.28, SD 3.6; F = 0.31, n.s.). On this measure also, the BB group failed to catch up on the early lead of the JP group, with the early lead in this case incontrovertibly attributable to the intervention.

(b) Phonic knowledge

As stated above, and shown in Table 4, the JP group was also significantly ahead of the BB group at pretest on all three tests of phonic knowledge. Despite this, they made significantly more progress than the BB group during the intervention, on both recognition and recall of letter sounds and on writing sounds to dictation. In the year following intervention, the BB group showed signs of catching up, making significantly more progress than the JP group on all three tests. However, they did not catch up on the early lead of the JP group: the effects of early intervention are still apparent at the end of the second year in school, with JP children significantly ahead of BB children at T3 on all three tests. Furthermore, the BB group's T3 scores were again statistically indistinguishable from the JP group's T2 scores (one year earlier) on both letter sound recognition (mean BB T3 = 22.2, SD 5.3; mean JP T2 = 23.2, SD 4.8; $F = 1.09$, n.s.) and writing sounds to dictation (mean BB T3 = 6.26; mean JP T2 = 6.59; $F = 0.58$, n.s.). On the harder measure of letter sound recall, which included knowledge of consonant and vowel digraphs, the BB group at T3 were performing significantly worse than the JP group one year earlier, at T2 (mean JP T2 = 28.44, SD 10.7; mean BB T3 = 22.64, SD 8.3; $F = 10.17, p < .002$). Thus, on all three phonics measures, the BB group failed to catch up on the early lead of the JP group, with at least part of the early lead in phonic knowledge definitely attributable to the intervention.

(c) Reading measures

As shown in Table 5, children were well matched across intervention groups at pretest on all four reading measures, with no significant pretest differences, and most children completely unable to read either words or nonwords. Differences in favour of the JP group in progress during the intervention just missed statistical significance on the two standardised tests. However, on the two experimental measures, Read words and Read nonwords, the JP group made significantly more progress during intervention than the BB group. Thus, the intervention did result in some immediate effects on reading. In the year following intervention, differences in favour of the JP group in reading

Table 4. Experimental measures 1: Mean differences in phoneme awareness and phonics knowledge over time

Measure	N items		(pretest) T1		Gains T1–T2		Gains T2–T3		(del. post-test) T3	
			JP	BB	JP	BB	JP	BB	JP	BB
Init. phon. ident'n	24	mean	12.59	6.82	7.98	5.74	2.6	6.9	22.82	19.57
		SD	(9.1)	(7.9)	(8.8)	(8.5)	(6.1)	(8.3)	(1.4)	(6.7)
			$z = 3.66$		$F = 1.8$		$F = 9.32$		$F = 12.31$	
			$p = .0002$		n.s.		$p < .003$		$p < .0007$	
Phoneme segment'n	11	mean	0.54	0.14	3.81	0.52	4.1	4.2	8.35	4.67
		SD	(1.9)	(0.6)	(3.3)	(1.5)	(3.1)	(3.5)	(2.8)	(3.7)
			$z = 1.13$		$z = 5.65$		$F = 0.01$		$F = 35.15$	
			n.s.		$p < .00001$		n.s.		$p < .00001$	
Letter Sound recog'n	26	mean	14.91	8.51	8.43	5.98	2.3	7.9	25.49	22.20
		SD	(6.9)	(7.3)	(5.5)	(6.7)	(4.2)	(6.4)	(1.2)	(5.3)
			$z = 4.36$		$F = 4.29$		$F = 28.50$		$F = 20.56$	
			$p < .00001$		$p < .05$		$p < .00001$		$p < .00001$	
Letter sound recall	43	mean	11.91	4.30	17.00	6.59	3.9	12.0	32.37	22.64
		SD	(8.8)	(6.7)	(9.2)	(6.7)	(7.0)	(7.3)	(7.9)	(8.3)
			$z = 4.48$		$F = 45.05$		$F = 34.74$		$F = 39.63$	
			$p < .00001$		$p < .00001$		$p < .00001$		$p < .00001$	
Write sounds	10	mean	2.98	1.35	3.75	2.22	0.5	1.5	8.22	6.26
		SD	(1.9)	(1.7)	(2.2)	(2.2)	(1.9)	(2.3)	(1.8)	(2.1)
			$z = 4.49$		$F = 12.95$		$F = 5.54$		$F = 28.67$	
			$p < .00001$		$p < .0005$		$p < .03$		$p < .0001$	

progress as measured on the BAS single word reading test were highly significant, suggesting that their word recognition skills had received a lasting boost from the early intervention, and, indeed, at the end of the second year in school the JP children were significantly ahead of the BB group on this test. There was no differential progress in the year following intervention when the measure of progress was performance on the Young's test. This is possibly because this is a forced choice test with chance therefore a factor: because different numbers of distractor items are available for different test stimuli, it is impossible to calculate, and correct scores for, chance. However, at T3, the JP group were also significantly ahead of the BB group on this reading measure too.

Table 5. Experimental measures 2: Mean differences in reading and spelling over time

Measure	*N* items		(pretest) T1		Gains T1–T2		Gains T2–T3		(del. post-test) T3	
			JP	BB	JP	BB	JP	BB	JP	BB
BAS reading (raw)		mean	1.51	0.25	7.15	2.81	27.5	20.5	35.85	22.78
		SD	(4.7)	(0.9)	(9.3)	(4.1)	(16.7)	(16.0)	(21.4)	(18.6)
			$z = 1.39$		$z = 1.92$		$F = 4.86$		$F = 11.79$	
			n.s.		n.s. ($p < .06$)		$p < .03$		$p < .0008$	
Young reading (raw)		mean	4.71	3.44	5.39	3.84	9.1	7.9	19.24	14.80
		SD	(4.9)	(4.4)	(4.1)	(4.5)	(4.9)	(4.8)	(7.2)	(7.1)
			$z = 1.35$		$F = 3.57$		$F = 1.80$		$F = 10.62$	
			n.s.		n.s. ($p < .07$)		n.s.		$p < .002$	
Read words	14	mean	1.26	0.61	4.43	2.69	6.3	7.2	11.85	10.30
		SD	(2.9)	(1.2)	(4.0)	(3.2)	(4.4)	(4.0)	(3.6)	(4.4)
			$z = 0.87$		$F = 6.28$		$F = 1.42$		$F = 4.18$	
			n.s.		$p < .02$		n.s.		$p < .05$	
Read CVC nonwords	10	mean	0.69	0.07	2.55	0.52	3.5	3.6	6.70	4.07
		SD	(2.2)	(0.4)	(2.9)	(1.6)	(2.9)	(3.6)	(3.2)	(3.9)
			$z = 1.85$		$z = 4.22$		$F = 0.06$		$F = 14.56$	
			n.s.		$p < .00001$		n.s.		$p < .0002$	
Write words	10	mean	0.30	0.09	3.00	0.59	3.8	4.4	7.07	4.89
		SD	(0.9)	(0.3)	(2.8)	(1.5)	(2.6)	(2.6)	(2.5)	(3.0)
			$z = 0.81$		$z = 5.12$		$F = 1.24$		$F = 17.04$	
			n.s.		$p < .00001$		n.s.		$p = .0001$	

Progress over time on the two non-standard measures (Read words, Read nonwords) happened as predicted, with significantly more progress in the JP group during intervention, equal progress in both groups in the year following intervention, and an absolute advantage for the JP group at the end of the second year in school. Thus, the early intervention can be seen to have exerted a lasting influence on progress in reading.

(d) Writing measures

Performance on Writing Words to Dictation also happened as predicted. There was no significant difference between the two groups at pretest. The JP group made significantly more progress during the intervention. There was no difference in progress in the year following intervention, yet at T3 the JP group were still significantly ahead of the BB group, indicating also a lasting effect from the early intervention on spelling skills.

(e) Additional measures given at delayed post-test

As shown in Table 6, on the Neale standardised reading test given at T3 at the end of the second year in school, the JP children scored significantly better for reading accuracy than the BB children, and differences in reading comprehension in favour of the JP group just failed to reach statistical significance.

Table 6. Mean differences on additional measures given at delayed post-test

Measure	*N* items		Jolly Phonics	Big Books	Test statistic	Significance level
Neale Accuracy raw score		mean (SD)	18.72 (14.78)	11.95 (12.60)	$F = 6.71$	$p = .01$
Neale Comprehension raw score		mean (SD)	5.02 (4.24)	3.54 (3.66)	$F = 3.77$	n.s. ($p < .06$)
Clay Dictation	37	mean (SD)	32.91 (4.14)	26.38 (10.84)	$F = 17.07$	$p < .0001$
Schonell spelling raw score		mean (SD)	17.69 (10.56)	9.54 (8.49)	$F = 20.15$	$p < .00001$
Schonell Spelling Age		mean (SD)	6.76 (1.07)	5.84 (1.00)	$F = 21.93$	$p < .00001$
BAS reading age		mean (SD)	7.09 (1.20)	6.28 (1.07)	$F = 13.76$	$p = .0003$

The JP group also scored significantly better on the Clay dictation test: of particular note here is the fact that no child in the JP group scored below 24/37 on this test, whereas the lowest score in the BB group was 0/27, and 12 children in the BB group scored below 24/37. Thus, in the JP group, the 'tail' of low achievement in writing was simply abolished.

On the Schonell spelling test, the JP group scored significantly better than the BB group: in terms of raw scores, they spelled almost twice as many words correctly as the BB group. In terms of spelling age, the JP group had a mean advantage of 11 months over the BB group.

Also shown in Table 6 are reading ages on the BAS test at T3, which show that in terms of reading age the JP group had a mean advantage of 10 months over the BB group.

Thus, the additional tests given at delayed post-test provide further evidence of the lasting effects of the early intervention, with children in the JP group outperforming children in the BB group by almost one year in both reading and spelling.

Discussion

The results presented above show strong, specific and significant positive effects of the JP intervention on the development of reading and writing. They show that the early intervention was responsible for the differences in reading, writing and phonic skills observed at final post-test, as progress in the year following intervention was similar in both groups. They provide very strong additional support for the view that early, structured, focused and rapid teaching of phoneme segmentation and blending skills and of grapheme-phoneme correspondences does accelerate development of these skills and acquisition of this knowledge in 5-year-olds. They extend this finding to children learning English as a second language who have initially very poor receptive vocabularies for English. They also provide further strong support for the view that phoneme awareness, segmentation and blending skills, and grapheme-phoneme correspondence knowledge influence the development of reading and writing skills, and that this leads to a lasting advantage for children who acquire these prerequisites at least as soon as (if not before) they are formally introduced to tuition in reading and writing. Moreover, most children can very rapidly acquire the concepts and knowledge taught, and can do so without the necessity for small-group teaching. The materials used in the present study were originally designed by a teacher for teachers to use, and our results additionally demonstrate that teachers need very little training or support to use these materials to good effect.

It remains to be seen whether the understanding and knowledge of how the alphabet functions to represent spoken language gained by children in the JP group in the 12 weeks intervention will be sufficient to enable them to develop self-teaching printed word recognition systems: a further follow-up is planned at the end of their third year in school, when this will be the major question to be investigated. We also intend to investigate whether an early advantage in reading skills has beneficial effects on oral vocabulary development: it is possible that the better progress in oral vocabulary development of the JP group during the year following intervention was not merely an inevitable catch-up on the BB group's earlier superiority, but was perhaps also due to their better reading ability. It will also be of interest to see whether the JP group retain their early advantage over the BB group in reading and spelling, and whether, as results from Byrne & Fielding-Barnsley (1995) give cause to hope, this later extends to significant differences also in reading comprehension.

Acknowledgments

Thanks are due to the London Docklands Development Corporation, the St Katharine and Shadwell Trust, The Mercers' Company, and Canary Wharf Ltd, who supplied the funding without which this research could not have been undertaken; to Liz Poole for supervising the early data collections, and to Liz Poole, Philippa Dodds, Lesley Zuke, Sandy Barker, Frances Bodger and Nicola Macfarlane for testing the children. Thanks are also due to the schools who agreed to take part in the research, and who remained committed to the principle of research even when the going got tough. Above all, thanks are due to the teachers who willingly adapted or changed their practice during the intervention period, and to the children who happily underwent repeated testing, even when asked to do rather difficult tasks.

References

Ball, E. W. & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, **26**, 49–66.

Blachman, B. A., Ball, E. W., Black, R. S. & Tangel, D. M. (1994). Kindergarten teachers develop phoneme awareness in low-income, inner-city classrooms: Does it make a difference? *Reading and Writing*, **6**, 1–18.

Bradley, L. & Bryant, P. (1983). Categorising sounds and learning to read: A causal connection. *Nature*, **30**, 419–421.

Brady, S., Fowler, A., Stone, B. & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia*, **44**, 26–58.

Bruck, M. & Genesee, F. (1995). Phonological awareness in young second language learners. *Journal of Child Language*, **22**, 307–324.

Bryant, P. (1998). Sensitivity to onset and rime does predict young children's reading: A comment on Muter, Hulme, Snowling & Taylor (1997). *Journal of Experimental Child Psychology*, **71**, 29–37.

Bryant, P., Bradley, L., Maclean, M. & Crosland, J. (1989). Nursery rhymes, phonological skills and reading. *Journal of Child Language*, **16**, 407–428.

Byrne, B. & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology*, **83**, 451–455.

Byrne, B. & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology*, **87**, 488–503.

Campbell, R. & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, **13**, 61–68.

Clay, M. (1979). *The Early Detection of Reading Difficulties.* London: Heinemann.

Cunningham, A. E. (1990). Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, **50**, 429–449.

Duncan, L., Seymour, P. H. K. & Hill, S. (1997). How important are rhyme and analogy in beginning reading? *Cognition*, **63**, 171–208.

Duncan, L., Seymour, P. H. K. & Hill, S. (In press). Small to large unit progression in metaphonological awareness and reading? *Quarterly Journal of Experimental Psychology*, Section A.

Dunn, L. M., Dunn, L. M., Whetton, C. & Burley. (1997). *The British Picture Vocabulary Scales*, 2nd ed. Windsor: NFER-Nelson.

Elliott, C. D., Murray, D. J. & Pearson, L. S. (1983). *British Ability Scales.* Windsor: NFER-Nelson.

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

Hatcher, P. J., Hulme, C. & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, **65**, 41–57.

Hoien, T., Lundberg, I., Stanovich, K. & Bjaalid, I. (1995). Components of phonological awareness. *Reading and Writing*, **7**, 171–188.

Holdaway, D. (1979). *The foundations of literacy.* Sydney, NSW: Ashton Scholastic.

Hulme, C., Muter, V. & Snowling, M. (1998). Segmentation does predict early progress in learning to read better than rhyme: A reply to Bryant. *Journal of Experimental Child Psychology*, **71**, 39–44.

Johnston, R. S. & Watson, J. (1997). Developing reading, spelling and phonemic awareness skills in primary school children. *Reading*, July, 37–40.

Lloyd, S. (1992). *The phonics handbook*. UK: Jolly Learning.

Lundberg, I., Frost, J. & Petersen, O-P. (1988). Effects of an intensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, **23**, 263–284.

Maclean, M., Bryant, P. & Bradley, L. (1987). Rhymes, nursery rhymes and reading in early childhood. *Merrill-Palmer Quarterly*, **33**, 255–281.

McGuiness, D., McGuiness, C. & Donohue, J. (1995). Phonological training and the alphabetic principle: Evidence for reciprocal causality. *Reading Research Quarterly*, **30**, 830–852.

Muter, V., Hulme, C., Snowling, M. J. & Taylor, S. (1997). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, **65**, 370–396.

Neale, M. D. (1997). *Neale Analysis of Reading Ability-Revised.* Windsor: NFER-Nelson.

Schonell, F. (1985). *Essentials in teaching and testing spelling*. Basingstoke: Macmillan.

Share, D. L. (1995). Phonological recoding and self-teaching: *Sine qua non* of reading acquisition. *Cognition*, **55**, 151–218.

Stuart, M. (1995). Prediction and qualitative assessment of five- and six-year-old children's reading: A longitudinal study. *British Journal of Educational Psychology*, **65**, 287–296.

Stuart, M. & Masterson, J. (1992). Patterns of reading and spelling in 10-year-old children related to prereading phonological abilities. *Journal of Experimental Child Psychology*, **54**, 168–187.

Stuart, M., Masterson, J., Dixon, M. & Quinlan, P (1999). Inferring sublexical correspondences from sight vocabulary: Evidence from 6- and 7-year olds. *Quarterly Journal of Experimental Psychology*, **52A**, 353–366.

The National Literacy Strategy: Framework for Teaching (1998) London: DFEE Standards and Effectiveness Unit.

Thompson, G. B., Cottrell, D. S. & Fletcher-Finn, C. M. (1996). Sublexical orthographic-phonological relations early in the acquisition of reading. *Journal of Experimental Child Psychology*, **62**, 190–222.

Watson, J. & Johnston, R. S. (1999). The effects of synthetic phonics teaching on reading, spelling and phoneme awareness skills. Submitted for publication.

Wiig, E. H., Secord, W. & Semel, E. (1992). *Clinical evaluation of language fundamentals-preschool*. New York: The Psychological Corporation.

Willows, D. (1996). *Jolly phonics*. Paper presented at a seminar, Institute of Education, University of London.

Yopp, H. K. (1988). The reliability and validity of phonemic awareness tests. *Reading Research Quarterly*, **23**, 159–177.

Young, D. (1985). *Group Reading Test*. London: Hodder & Stoughton Educational.

Zinna, D. R., Liberman, I. Y. & Shankweiler, D. (1986). Children's sensitivity to factors influencing vowel reading. *Reading Research Quarterly*, **21**, 465–480.

Received 24 November 1998; revised version received 10 March 1999