Improving early language and literacy skills: differential effects of an oral language versus a phonology with reading intervention

Claudine Bowyer-Crane,¹ Margaret J. Snowling,¹ Fiona J. Duff,¹ Elizabeth Fieldsend,¹ Julia M. Carroll,² Jeremy Miles,³ Kristina Götz¹ and Charles Hulme¹

¹Department of Psychology, University of York; ²Department of Psychology, University of Warwick; ³Department of Health Sciences, University of York

Background: This study compares the efficacy of two school-based intervention programmes (Phonology with Reading (P + R) and Oral Language (OL)) for children with poor oral language at school entry. Methods: Following screening of 960 children, 152 children (mean age 4;09) were selected from 19 schools on the basis of poor vocabulary and verbal reasoning skills and randomly allocated to either the P + R programme or the OL programme. Both groups of children received 20 weeks of daily intervention alternating between small group and individual sessions, delivered by trained teaching assistants. Children in the P + R group received training in letter-sound knowledge, phonological awareness and book level reading skills. Children in the OL group received instruction in vocabulary, comprehension, inference generation and narrative skills. The children’s progress was monitored at four time points: pre-, mid- and post-intervention, and after a 5-month delay, using measures of literacy, language and phonological awareness. Results: The data are clustered (children within schools) and robust confidence intervals are reported. At the end of the 20-week intervention programme, children in the P + R group showed an advantage over the OL group on literacy and phonological measures, while children in the OL group showed an advantage over the P + R group on measures of vocabulary and grammatical skills. These gains were maintained over a 5-month period. Conclusions: Intervention programmes designed to develop oral language skills can be delivered successfully by trained teaching assistants to children at school entry. Training using P + R fostered decoding ability whereas the OL programme improved vocabulary and grammatical skills that are foundations for reading comprehension. However, at the end of the intervention, more than 50% of at-risk children remain in need of literacy support. Keywords: Early intervention, oral language, phonological awareness, early literacy, RCT.

It is well established that phonological skills are fundamental to alphabetic literacy (Goswami & Bryant, 1990; Byrne, 1998). However, aspects of oral language ability beyond phonology provide the foundation for reading comprehension (Oakhill, Cain, & Bryant, 2003; Muter, Hulme, Snowling, & Stevenson, 2004). The Simple View of Reading (Gough & Tunmer, 1986) captures this differential role of different language skills in reading development. Within this framework, reading comprehension depends upon the interaction of decoding skill and linguistic comprehension. Individual differences in decoding ability are predicted by variations in letter knowledge and phoneme awareness (Bowey, 2005, for a review). In contrast, variations in linguistic comprehension depend upon a range of factors including vocabulary and grammatical abilities and resources such as attention. A similar distinction has been drawn by Whitehurst and Lonigan (1998) who differentiated two domains of emergent literacy: ‘inside-out’ skills (e.g., letter knowledge and phoneme awareness) and ‘outside-in’ skills (e.g., vocabulary and grammatical skills).

Building on the Simple View, Bishop and Snowling (2004) proposed a two-dimensional model of reading impairment with phonological skills lying on one dimension, and non-phonological skills (e.g., semantics and syntax) lying on the other. According to this model, the risk of word-level decoding difficulties in reading is carried by phonological deficits, whereas the risk of reading comprehension difficulties is associated with deficits in non-phonological language skills. Thus, there is a strong theoretical rationale for early years teaching to foster good speaking and listening skills (Rose, 2006) and, more specifically, for intervention programmes to target oral language skills in language-delayed children who are likely to be at risk of literacy problems.

The majority of research on reading intervention has been concerned with ameliorating word-level reading difficulties (Torgesen, 2005; Troia, 1999). Evidence indicates that interventions combining phonological training with reading are successful in facilitating reading development in poor readers (Gillon, 2000, 2002; Hatcher et al., 2006a; Hatcher,
Trials Unit to one of two 20-week intervention programmes: a Phonology with Reading Intervention (P + R), or an Oral Language Intervention (OL). Assessments were made at pre-test (t1), mid-test after 10 weeks (t2), post-test at the end of the 20-week intervention (t3) and five months after the intervention had ceased (t4). At t4, a sample of 564 children, drawn from 18 of the schools originally screened, were assessed on tests of single word reading. Ethical approval was obtained from the Ethics Committee, Department of Psychology, University of York and all assessments were carried out with informed consent from headteachers and parents where appropriate.

Participants
Details of the recruitment, selection, and allocation of the participants are summarised in Figure 1, in accordance with the CONSORT statement (Moher, Schulz, & Altman, 2001). Twenty-three mainstream schools were involved at the outset of the study. From these schools, every child was screened at school entry, in autumn 2004 (in the UK children enter school in the academic year in which they turn five).

Following screening, three schools were deemed unsuitable for continued involvement given the relatively high performance of their children on our language measures. In each of the remaining 20 schools, the 10 children with the lowest age-standardised scores on the Picture Naming sub-test from the Wechsler Pre-School and Primary Scale of Intelligence-III UK (WPPSI-III UK; Wechsler, 2003) were selected as possible candidates for intervention. To validate this initial selection, more extensive individual assessments were conducted with each of these children using further language measures, and each child’s WPPSI Vocabulary and Word Reasoning scaled scores were averaged to form a verbal composite measure. The 8 children with the lowest scores on this verbal composite measure in each school were selected to receive intervention. Cut-points varied by school and ranged from a mean scaled score of 5.56 to 8.75. One school then withdrew, leaving a total of 152 children in 19 schools. The 8 children in each school were randomly allocated without restriction to the P + R or OL Intervention (Table 1). Gender was equivalent across groups (P + R = 52.6% male, OL = 47.4% male). In addition, the participants were rated on the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997): significant behavioural problems were reported for 21.6% of the P + R group and 22.4% of the OL group. Attrition rates differed between groups; 9 children were lost from the P + R group compared to 1 child from the OL group.

Tests and procedures
Testing was carried out by two members of the research team (CC, FD) with assistance from students from the Department of Psychology, University of York when necessary. All testers were trained in the administration of the tests prior to delivering the assessments. To ensure that investigators were blind to group membership when

Method
We conducted a randomised controlled trial (RCT) in which children were randomly allocated by the York

Claudine Bowyer-Crane et al.
December 2004: $N = 960$ children in 23 schools screened for participation in early intervention

3 schools excluded ($n = 100$): performance too high

January 2005: $n = 200$ children selected from remaining 20 schools ($n = 861$) for further testing

$= 4$ children unavailable for testing
$n = 196$ children tested

Allocation
Feb 2005: $n = 160$ children selected to take part in the intervention programmes ($n = 8$ from each school) and randomly allocated to one arm of the intervention project.

1 school withdrawn from programme
$n = 8$ (4 children from each arm)
$n = 17$ children replaced following discussion with teacher


**Figure 1** Flowchart showing selection and attrition of participants in accordance with the Consort guidelines

testing, the research team was not involved in the allocation of children to the interventions. In the event of schools needing to contact the research team to talk about the programme, each school was allocated one member of the team as their contact. The other member of the team was assigned to carry out the assessments in that school. As far as was practicable, tests were presented in a fixed order to all participants. All tests had good reliability (ranging from .84 to .98).

**Screening phase**

The screening battery consisted of an expressive language test and a test of nonword repetition. In addition, short assessments of letter knowledge, reading, and writing were given but data were not analysed because of floor effects.

**Nonword repetition**

30 items from the Children’s Test of Nonword Repetition (Gathercole & Baddeley, 1996) were administered. Following piloting, the full version of this test was judged to be too demanding for children of this age.

**Picture Naming** (also given at t4) from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) was administered as a test of expressive vocabulary.

**Pre-, mid-, post- and maintenance tests**

Primary outcomes were word-level reading skills, vocabulary and grammar. The assessment measures used for the pre-, mid-, post- and maintenance test phases were selected to assess children’s progress in the intervention towards these outcomes, and included...
Table 1 Characteristics of children in each arm of the intervention at t0 (screening): Mean scores with standard deviations in parentheses

<table>
<thead>
<tr>
<th></th>
<th>Phonology with Reading</th>
<th>Oral Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender M:F</td>
<td>40.36; 36.40</td>
<td></td>
</tr>
<tr>
<td>Age (months)*</td>
<td>57.53 (3.51); 56.83 (3.22)</td>
<td></td>
</tr>
<tr>
<td>Letter Knowledge (max 14)</td>
<td>4.34 (3.64); 3.78 (3.16)</td>
<td></td>
</tr>
<tr>
<td>Early Word Recognition (max 6)</td>
<td>0.83 (0.84); 0.76 (0.87)</td>
<td></td>
</tr>
<tr>
<td>Non-word repetition (max 30)</td>
<td>12.30 (7.21); 13.58 (6.82)</td>
<td></td>
</tr>
<tr>
<td>WPPSI-III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Naming</td>
<td>6.01 (1.49); 6.33 (1.25)</td>
<td></td>
</tr>
<tr>
<td>Word Reasoning</td>
<td>7.67 (2.54); 7.65 (2.09)</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>6.21 (2.13); 6.67 (2.26)</td>
<td></td>
</tr>
<tr>
<td>Block Design</td>
<td>6.92 (2.98); 6.86 (3.23)</td>
<td></td>
</tr>
<tr>
<td>SDQ Total Deviance score (max = 20)</td>
<td>9.45 (6.12); 10.54 (6.43)</td>
<td></td>
</tr>
<tr>
<td>Free school meals</td>
<td>28.9% (18.4%)</td>
<td></td>
</tr>
</tbody>
</table>

*Age months = mean age of children collapsed across screening and administration of cognitive battery; *scaled scores where population mean = 10; SD = 15; P + R group n = 65; OL group n = 68.

measures tapping skills directly targeted by the interventions. Owing to limited access to children during the school day, and to avoid placing unreasonable demands on them, not all tests were given on each occasion. Brief details of the test battery are provided below, grouped according to construct (full details at http://www.york.ac.uk/res/crl/crl_Nuffield.html). At least one measure of each construct was given at each occasion. Brief details of the test battery are provided below, grouped according to construct (full details at http://www.york.ac.uk/res/crl/crl_Nuffield.html). At least one measure of each construct was given at each occasion, except at t4 when phonological skills were not directly assessed, though indirect tests tapping phonological reading and spelling strategies were given.

General cognitive ability was assessed at t1 with WPPSI-II<sup>x</sup> Block Design, Vocabulary, Word Reasoning. The Matrix Reasoning scale was given at t4.

Phonological measures

Phoneme awareness (t1, t2, t3) was assessed with the initial phoneme detection component of the Sound Isolation Task (Hulme, Caravolas, Målkova, & Briggstocke, 2005).

Phoneme Completion (t3) from the Phonological Abilities Test (PAT; Muter, Hulme, & Snowling, 1997) also measured phoneme awareness.

Phoneme Blending, Segmentation and Deletion (t3) tasks from the Test of Phonological Awareness (Hatcher, 2000) were used to assess children's ability to segment and blend words.

Language measures

Expressive grammar (t1, t3, t4) was assessed using The Action Picture Test (Renfrew, 2003).

Narrative skill (t1, t2, t3) was measured using The Bus Story (Renfrew, 1991). According to the manual, this test measures ‘the ability to give a coherent description of a continuous series of events’. A score was calculated for both information given and average sentence length.

Specific vocabulary (t2, t3, t4): knowledge of words taught directly in the OL programme was measured using a combination of picture naming and questions requiring one-word answers (e.g., what is the opposite of back?) (maximum = 25).

Listening comprehension (t1, t3) was assessed using recordings of stories taken from the Neale Analysis of Reading Ability II (NARA II; Neale, 1997) (Levels 1 and 2; Form 2). A point was awarded for each correct answer (maximum = 12).

Literacy measures

Letter identification (t1, t2, t3, t4): Children were asked to identify by sound 24 of the 26 letters in the English alphabet (t1 and t2) or all 26 letters (t3 and t2).

Single word reading (t1, t2, t3, t4): Single word reading ability was assessed using the Early Word Recognition Test (EWR; Hatcher et al., 1994), with testing being discontinued after 5 consecutive errors. The British Ability Scales II (BAS II; Elliott, Smith, & McCulloch, 1997) Word Reading scale was given to those children who read 30 or more items on the EWR test.

Reading comprehension (t3): Children read two short stories: the Level 1 passage taken from Form 1 of the NARA II, and Passage 1, Form 1 from the Gray Oral Reading Tests 4 (GORT 4; Wiederhold & Bryant, 2001). A point was awarded for each correct answer (maximum = 9).

Prose reading accuracy (t3): A measure of reading accuracy was taken during administration of the reading comprehension test (maximum = 46).

Nonword Reading (t4): The Graded Nonword Reading Test (GNWRT; Snowling, Stothead, & McLean, 1996) was given at t4 as a measure of decoding.

Spelling (t1, t2, t3, t4): Five words were presented as pictures to be named and spelled. They were scored for items correct and percentage consonants correct. At t4, five more complex items were added to the spelling test.

Behavioural assessment

Strengths and Difficulties Questionnaire (t1, t3) (Goodman, 1997) was completed by class teachers and teaching assistants for each child.

Socioeconomic status

Post-codes for 133 children were obtained from their schools and used to derive an estimate of socioeconomic deprivation: [http://www.neighbourhood.gov.uk/page.asp?id=1057]. We also obtained data from schools on whether each child was in receipt of free school meals.

Intervention programmes

The two intervention programmes shared the same structure and were designed to run over two 10-week periods. Children received alternating daily one-to-one (20 minute) and group (30 minute) lessons. A manual was written for each programme documenting activities and procedures. Each 10-week period was divided into an initial introduction week followed by three 3-week
teaching blocks, each further divided into a 2-week instruction period and a third consolidation week. Each programme was run to a fixed time schedule (see Table 2).

**Phonology with Reading (P + R) programme**

The P + R programme had three main components: letter-sound knowledge, phonological awareness (including articulatory awareness) and reading books at the instructional level. Direct teaching in sight word recognition was also included in order to build up children’s reading vocabulary.

Children were trained in letter-sound knowledge using the Jolly Phonics programme (Lloyd, 1998). Letter-sound knowledge was reinforced through reading, writing, and phonological awareness activities including blending and segmenting.

Phoneme awareness was taught for approximately 5 minutes in each session through blending and segmenting exercises in line with guidelines provided by the National Reading Panel (NICHD, 2000) on phonemic awareness instruction. This work was done using multi-sensory techniques using a scaffolding approach to ensure that children were working at a suitable level. For those children not yet able to pronounce specific phonemes, work on articulatory awareness and phoneme production was included in the individual sessions.

Children interacted with books on a regular basis and were encouraged to link letter-sound knowledge and phoneme awareness in the context of listening to storybooks (Ukrainetz, Cooney, Dyer, Kysar, & Harris, 2000). In each individual teaching session, the child read two books to the teaching assistant (TA) who took a running record when the child read the first book in order to assess the level at which the child was reading (Hatcher, 2000). The TA then introduced a new book, which the child read alone first, and then again with the TA to encourage fluency. In the case of children who could not yet read at all, the ‘cut-up’ story activity was substituted (Clay 1985; Hatcher et al., 2006b).

**Oral Language programme**

The Oral Language (OL) programme included direct instruction to develop vocabulary, inferencing, expressive language and listening skills. Activities were adapted from a number of sources, including Rhodes to Language (Rhodes, 2001), Time to Talk (Schroeder, 2001), and materials from Black Sheep Press (e.g., Rippon, 2002).

Since listening skills are fundamental to language development, specifically targeted activities required children to listen to and retain information in order to complete a task. Vocabulary to be taught was selected according to two criteria; (i) that it was age-appropriate and instructional, and (ii) that it was related to one of the selected topics. The vocabulary to be taught included a selection of nouns, verbs, comparatives and spatial terms, as well as question words. All words were taught using methods that encouraged children to use them in different contexts (Beck, McKeown, & Kucan, 2002). New vocabulary was introduced every group session, and reinforced in the following group session and in individual sessions.

Narrative work was included to encourage expressive language development and grammatical competence. In the group sessions, many activities revolved around creating stories (e.g., ‘washing line’ activity from Time to Talk; Schroeder, 2001). A specially designed narrative task in which children told a story from cartoon sequences was used in individual sessions. TAs transcribed these narratives and used them as a basis for elaborating the story in the next session.

Independent speaking was encouraged in all sessions through the interactive nature of the programme. Specific activities included ‘show and tell’ sessions, and ‘magic sack’ activities (describing an object to the rest of the group). TAs were taught to monitor children’s grammatical errors and to model the correct forms when errors occurred. Question words were taught throughout the programme and, as well as answering questions, children were encouraged to seek information by generating their own questions.

Teaching assistants were nominated by their schools; they received 4 days’ training before the intervention began and one day mid-way through. In addition, they were supported in fortnightly group tutorials by the research team and observed once teaching to assess treatment fidelity, when they also received feedback.

**Results**

We wished to compare the mean difference between groups on several outcome variables at the end of the intervention (t3) and six months after the intervention had finished (t4). The data from the 152 participants were clustered within 19 schools; within each school the same teaching assistant taught in both arms. The data were therefore analysed using complex samples analyses (SPSS 15.0) giving robust standard errors that take account of the non-independence of observations within clusters. Group differences at t3, at the end of the intervention, and at t4, six months after the intervention, were assessed using dummy coding of group in a regression model after controlling for differences associated with gender, chronological age and pre-intervention levels of performance on the same task (the autoregressor) when this was available. These analyses are equivalent to performing analyses of covariance controlling for gender, age and the autoregressor.

Table 3 shows the means and standard deviations for all measures at times 1, 2, 3 and 4 grouped according to intervention programme. Cell sizes (t1, t2, t3, t4) varied from 67 to 76 owing to variations in pupil attendance and cooperation.

For ease of interpretation Figures 2a and 2b show the differences between the two groups in z-score units. A difference of 1.0 in these figures represents a difference of 1 SD between the groups (equivalent to an effect size (Cohen’s d) of 1.0). The error bars represent robust 95% confidence intervals (CIs).
| **Table 2** Group and individual session schedule for P + R and OL programmes |
|-------------------------------|---------------------------------|---------------------|
| **Phonology with Reading**    | **Individual session (20 minutes)** | **Oral Language** |
| **Group session (30 minutes)**| **Revision**                     | **Group session (30 minutes)** |
| Ice Breaker                   | – Revision of letter             | – Introduce day of the week |
| – 4 minutes                   | – Revise/reinforce sounds learnt so far | – Talk about best listener |
| – 2 minutes                   | – 2 minutes                      | – 2 minutes          |
| Sound of the day              | Phonological awareness           | **Individual session (20 minutes)** |
| – Story and action            | – Segmenting or blending activities | **Vocabulary revision** |
| – Writing letter              | AND                             | – Revise vocabulary from last group session |
| – 8 minutes                   | Making sounds                   | – Concentrate on difficult vocabulary |
| – sound production exercises  |                                  | – 5 minutes          |
| if necessary                  |                                  |                      |
| – 3 minutes                   |                                  |                      |
| Book work                     | Sight word work                 | **Reinforcement**    |
| – Book for the sound of the day | Work on sight word vocabulary | – Reinforce vocabulary introduced |
| – Teacher-led ‘shared’ reading| – 5 minutes                     | in last session      |
| – 8 minutes                   |                                  | – 7 minutes          |
| Phonological awareness        | **Plenary**                     | **Narrative task**   |
| – Segmenting or               | Reading books                   | – Child to produce story to sequence of pictures |
| blending activities           | – Re-reading a book at the      | – Revisit story targeting areas for improvement |
| – 5 minutes                   | instructional level             | – 5 minutes          |
| Plenary                       | – Reading a new book            | **Speaking/listening/inferencing** |
| – Revise sound of the day     | – 10 minutes                    | – One or two activities to target specific skills |
| – Word sticker                |                                  | – 5 minutes          |
| – 5 minutes                   |                                  |                      |

Claudine Bowyer-Crane et al. / C211 2007 The Authors Journal compilation / C211 2007 Association for Child and Adolescent Mental Health.
Table 3  Data from main variables at beginning (t1) and end (t3) of intervention, according to intervention programme*(t2 & t4 scores given if measure not tested at these points). Raw mean scores and standard deviations in parentheses.

<table>
<thead>
<tr>
<th>Measure</th>
<th>P + R programme</th>
<th>OL programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter identification t1</td>
<td>13.69 (6.72)</td>
<td>14.12 (6.06)</td>
</tr>
<tr>
<td>Letter identification t3</td>
<td>23.92 (3.14)</td>
<td>22.19 (4.82)</td>
</tr>
<tr>
<td>EWR t1</td>
<td>4.88 (7.0)</td>
<td>3.04 (3.55)</td>
</tr>
<tr>
<td>EWR t3</td>
<td>21.08 (12.71)</td>
<td>16.27 (9.33)</td>
</tr>
<tr>
<td>Spelling t1</td>
<td>.18 (.58)</td>
<td>.08 (.32)</td>
</tr>
<tr>
<td>Spelling t3</td>
<td>1.47 (1.27)</td>
<td>.91 (.94)</td>
</tr>
<tr>
<td>% Consonants correct t1</td>
<td>20.38 (25.77)</td>
<td>19.41 (21.96)</td>
</tr>
<tr>
<td>% Consonants correct t3</td>
<td>62.19 (28.85)</td>
<td>55.41 (25.49)</td>
</tr>
<tr>
<td>Nonword reading t4</td>
<td>4.26 (5.71)</td>
<td>2.03 (4.07)</td>
</tr>
<tr>
<td>Prose reading accuracy t3</td>
<td>28.45 (13.02)</td>
<td>23.28 (10.16)</td>
</tr>
<tr>
<td>Read comprehension t3</td>
<td>5.11 (1.86)</td>
<td>4.72 (1.54)</td>
</tr>
<tr>
<td>Phonological awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoneme awareness t1</td>
<td>7.49 (7.81)</td>
<td>8.30 (7.72)</td>
</tr>
<tr>
<td>Phoneme awareness t3</td>
<td>20.19 (8.38)</td>
<td>19.31 (8.36)</td>
</tr>
<tr>
<td>Phoneme seg/blend/del t3</td>
<td>7.48 (4.79)</td>
<td>4.05 (3.33)</td>
</tr>
<tr>
<td>Phoneme completion t3</td>
<td>5.50 (2.93)</td>
<td>4.89 (2.68)</td>
</tr>
<tr>
<td>Language measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture naming raw t0</td>
<td>13.82 (2.87)</td>
<td>14.30 (2.51)</td>
</tr>
<tr>
<td>Picture naming raw t4</td>
<td>20.36 (2.51)</td>
<td>20.17 (2.41)</td>
</tr>
<tr>
<td>Specific vocabulary t3</td>
<td>11.04 (3.14)</td>
<td>14.89 (3.47)</td>
</tr>
<tr>
<td>Expressive grammar t1</td>
<td>15.75 (6.05)</td>
<td>16.70 (5.47)</td>
</tr>
<tr>
<td>Expressive grammar t3</td>
<td>20.71 (5.12)</td>
<td>22.67 (5.13)</td>
</tr>
<tr>
<td>Sentence length t1</td>
<td>6.79 (2.27)</td>
<td>7.19 (2.01)</td>
</tr>
<tr>
<td>Sentence length t3</td>
<td>8.48 (2.59)</td>
<td>9.19 (2.19)</td>
</tr>
<tr>
<td>Narrative skill t1</td>
<td>12.44 (6.99)</td>
<td>13.69 (6.32)</td>
</tr>
<tr>
<td>Narrative skill t3</td>
<td>19.09 (7.41)</td>
<td>20.84 (7.29)</td>
</tr>
<tr>
<td>Listening comprehension t1</td>
<td>1.63 (1.40)</td>
<td>1.70 (1.36)</td>
</tr>
<tr>
<td>Listening comprehension t3</td>
<td>2.44 (1.63)</td>
<td>2.63 (1.84)</td>
</tr>
</tbody>
</table>

Note. EWR = early word recognition; Phoneme seg/blend/del = phoneme segmenting, blending and deletion; Sentence length was calculated from the narratives produced in the narrative skill task. (a) Complete data set available from author.**

These analyses were repeated, controlling for social class and an index of behaviour (SDQ total deviance) in separate analyses. When controlling for behaviour (in addition to the other covariates) the group difference in Early Word Reading at t3 was significant (difference = .262, 95% CI: lower = .008, higher = .517, p < .05) but the significance levels of other measures remained unchanged. When social class (in addition to the other covariates) was controlled, the group difference in phoneme awareness at t3 was significant (difference = .284, 95% CI: lower = .053, higher = .514, p < .05), while the difference between groups in grammatical skill at t3 was no longer significant (difference = .297, 95% CI: lower = −.649, higher = .054, p < .092).

In the absence of an untreated control group it is difficult to gauge the absolute impact of each intervention on literacy and language skills. To provide an estimate of how much each intervention fostered the children’s literacy development, their single word reading performance at t4 was compared with that of a large sample of their peers from the same classes (n = 564). A composite score derived from performance on the Early Word Reading and BAS Single Word Reading tests was used as a standard (with mean of 100, SD = 15) against which to describe the performance of the children from the intervention groups.

A standard score below 85 for reading (−1SD) was used to classify children as being ‘at risk’ of literacy difficulties. At the end of the intervention, 68.1% of the OL group remained at risk on this criterion compared with only 50% of the P + R group. Moreover, 7.1% of children in the P + R group now had above-average reading scores (greater than 115), while none of the OL children had scores in this range.

Discussion

This RCT compared two contrasting intervention programmes; one targeting phonological and early reading skills and the other targeting oral language skills. The results show that these programmes were effective in promoting different aspects of literacy and spoken language in children with poor oral language skills at school entry. As expected, the children who received the P + R programme made better progress in literacy skills and in phoneme awareness but somewhat surprisingly not in single word reading where the two groups were comparable. In contrast, the children who received the OL programme made better progress in vocabulary and grammatical skills. The effects obtained were moderate to large in size and were mostly maintained at follow-up 5 months after the intervention ceased, except for phonetic spelling where group differences evened out (although differences in raw spelling

© 2007 The Authors
Journal compilation © 2007 Association for Child and Adolescent Mental Health.
A particularly encouraging finding was that gains in reading and phonological skills generalised to the reading of novel words at t4. The failure to find a significant differential effect on certain key measures is disappointing at first sight. However, there are likely to be a number of explanations for these null effects, not least differences in test sensitivity. For example, benefits accrued to the P + R group in phoneme awareness as measured by a test requiring segmentation, blending and deletion but not in tests requiring phoneme isolation (a very difficult test) or phoneme completion (a relatively easy test). In a similar vein, the listening comprehension test suffered from floor effects, being more difficult than anticipated for children of this age group.

Given that much of the variance in reading comprehension in the early school years can be attributed to decoding abilities (Whitehurst & Lonigan, 1998), it was not surprising that the effects of training on this aspect of reading were weak.

**Figure 2** (a) Relative advantage of P + R group on reading and phonology measures at t3 and t4 expressed as z-score units with 95% Confidence Intervals. (Letter Id = letter identification, EWR = early word recognition, Read Comp = reading comprehension, Spell_PCC = spelling percentage consonants correct, Nonwd Read = nonword reading, Prose Read Acc = prose reading accuracy, Phon Aware = phoneme awareness (sound isolation), Phoneme Seg/Blend/Del = phoneme segmenting, blending and deletion, Phon Comp = phoneme completion. Significant differences marked ***p < .001, **p < .01, *p < .05. † No autoregressor available. Group, Gender and Age in months at t0 were entered into each analysis as covariates. (b) Relative advantage of OL group on language measures at t3 and t4 expressed as z-score units with 95% Confidence Intervals. (Specific Vocab = specific vocabulary, Pic Naming = picture naming, Exp Grammar = expressive grammar, List Comp = listening comprehension). Significant differences marked ***p < .001, **p < .01, *p < .05. † No autoregressor available. Group, Gender and Age in months at t0 were entered into each analysis as covariates.
Nonetheless, the effect of the OL programme on vocabulary and grammatical skills is noteworthy since such skills are reliable predictors of reading comprehension (e.g., Muter et al., 2004). The teaching programme included narrative training, which encouraged children to increase the length of sentences they used and to improve their use of the correct grammatical forms. In addition, one of the teaching principles embodied in the programme was the use of modelling by the teaching assistants. Thus, when a child produced an immature grammatical form the teaching assistant would model a more appropriate version of the sentence. It is possible that these strategies were in some way instrumental in helping to bring about a change in the children’s grammatical usage but this aspect of the programme was not directly evaluated in the current study.

The finding that the strongest training effects for the Phonology with Reading programme in the present study were for phonological awareness is consistent with previous research, although a direct comparison of effect sizes is hampered by the use of different measures and the conservative design of the present study which reported differential gains (in relation to treated controls). In a meta-analysis of the effects of phonological awareness training on reading, Bus and van IJzendoorn (1999) demonstrated effect sizes between $d = 0.01$ and $d = 0.20$ for phonological awareness, with weaker effect sizes in randomised and matched designs ($d = 0.76$). Our finding of an effect size of $d = 0.72$ for phoneme segmentation, blending and deletion is in line with this finding. Bus and van IJzendoorn report a greater range of effect sizes for reading outcomes ($d = -0.12$ to $d = 0.76$), with an overall effect size of $d = 0.44$; as in the present study, training effects were stronger on tests of pure decoding ($d = 0.85$; present study $d = 0.41$) than single word reading ($d = 0.34$; present study $d = 0.23$).

There is less evidence regarding the effectiveness of oral language interventions. The present gains in instructed vocabulary are in line with findings reported by Beck and McKeown (2007) Study 1) for rich vocabulary instruction of mean gains of approximately 3 words for children in grade 1 ($d = 0.74$), and approximately 5 words for children in kindergarten ($d = 1.17$) over a 10-week period. In the present study the mean gain for the Oral Language group in instructed vocabulary at t3 was approximately 5.23 words ($d = 1.02$). A recent meta-analysis of interventions for children with speech and language difficulties (Law, Garrett, & Nye, 2004) reported effect sizes ranging from .28 to 1.02 for expressive syntax following clinician-led speech and language therapy of more than 8 weeks. However, the present gains in grammar and narrative ability are more directly comparable with those of Davies, Shanks, and Davies (2004), who found significant improvements in both Action Picture Test (Grammar) and Bus Story (Information scores) following a narrative intervention programme run over the course of a school term and delivered by trained learning support assistants (effect sizes adjusted for maturation $d = .74$ and $d = .44$ respectively). The present study reported effect sizes at t3 of $d = .33$ for Action Picture Test Grammar and $d = .15$ for Bus Story Information. These effects are weaker than the effect sizes reported by Davies et al. (2004) but it must be borne in mind that the current intervention included narrative work only as a component, whereas Davies and colleagues focused exclusively on narrative.

The present findings extend previous research by showing that focused language intervention programmes can be delivered successfully by trained teaching assistants to 4- and 5-year-old children at risk of literacy difficulties. A programme fostering phonological skills and letter knowledge had a positive effect on children’s emergent reading (word recognition) skills, whereas a programme focusing on oral language improved aspects of receptive and expressive language, particularly vocabulary and grammatical skills. Our findings suggest that both programmes are valuable interventions for children during the early school years and although we have no direct evidence for this supposition, a promising approach would appear to be to provide children with training in the Oral Language programme before school entry to reduce the numbers of children at risk of reading difficulties, and to provide those with continuing difficulties with an integrated approach combining aspects of the P + R and OL programmes. Further research is required to investigate this hypothesis.

Author note
Jeremy Miles is now at the RAND Corporation, Santa Monica, California.

Acknowledgements
This research was funded by the Nuffield Foundation and supported by North Yorkshire County Council and a British Academy Research Readership to Margaret Snowling. We thank Jolly Phonics and Black Sheep Press for donating materials, and Naomi Meredith, Nicky Vowles, Rachel Harlow, Debbie Gooch, Ros Francis, Dimitra Ionnou, Lisa Henderson, Lizzie Bowen, Natalie Falkinder, Sarah Edwards, Emma Truelove, Kim Manderson, Jodie Unau, Michelle Cargan, Pam Baylis, Rachael McCool, Elisa Romeo, Meesha Warmington, Poppy Nash and Janet Hatcher for their assistance throughout this project. Finally we would like to thank the pupils, teaching assistants and schools who participated.
Correspondence to
Claudine Crane, Dept of Psychology, University of York, Heslington, York, YO10 5DD, UK; Email: c.craneforpsych.york.ac.uk

References


Manuscript accepted 5 September 2007
Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

<table>
<thead>
<tr>
<th>Instruction to printer</th>
<th>Textual mark</th>
<th>Marginal mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leave unchanged</td>
<td>⋯ ⋯ under matter to remain</td>
<td>1</td>
</tr>
</tbody>
</table>
| Insert in text the matter indicated in the margin | / through single character, rule or underline or through all characters to be deleted | / or /$
| Delete                 | / through letter or through characters | / or /$
| Substitute character or substitute part of one or more word(s) | / through character or where required | / or /
| Change to italics      | — under matter to be changed | ʃ
| Change to capitals     | — under matter to be changed | ʃ
| Change to small capitals | — under matter to be changed | ʃ
| Change to bold type    | ≈ under matter to be changed | ≈
| Change to bold italic  | ≈ under matter to be changed | ≈
| Change to lower case   | Encircle matter to be changed | ≈
| Change italic to upright type | (As above) | ≈
| Change bold to non-bold type | (As above) | ≈
| Insert ‘superior’ character | / through character or where required | / or /
| Insert ‘inferior’ character | (As above) | / or /
| Insert full stop       | (As above) | /
| Insert comma           | (As above) | /
| Insert single quotation marks | (As above) | /
| Insert double quotation marks | (As above) | /
| Insert hyphen          | (As above) | /
| Start new paragraph    | (As above) | /
| No new paragraph       | (As above) | /
| Transpose              | (As above) | /
| Close up               | linking characters | /
| Insert or substitute space between characters or words | / through character or where required | /
| Reduce space between characters or words | between characters or words affected | /