This is a copy of the final published version of a paper published via gold open access in *Journal of Child Psychology and Psychiatry*

This open access article is distributed under the terms of the Creative Commons Attribution Licence (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/88748

---

**Published paper**

10.1111/jcpp.12010
Efficacy of language intervention in the early years

Silke Fricke,1 Claudine Bowyer-Crane,2 Allyson J. Haley,3 Charles Hulme,4 and Margaret J. Snowling3

1Department of Human Communication Sciences, University of Sheffield, Sheffield; 2Department of Psychology, Sociology and Politics, Sheffield Hallam University, Sheffield; 3Department of Psychology, University of York, York; 4Division of Psychology and Language Sciences, University College London, London, UK

Background: Oral language skills in the preschool and early school years are critical to educational success and provide the foundations for the later development of reading comprehension. Methods: In a randomized controlled trial, 180 children from 15 UK nursery schools (n = 12 from each setting; M_age = 4;0) were randomly allocated to receive a 30-week oral language intervention or to a waiting control group. Children in the intervention group received 30 weeks of oral language intervention, beginning in nursery (preschool), in three group sessions per week, continuing with daily sessions on transition to Reception class (pre-Year 1). The intervention was delivered by nursery staff and teaching assistants trained and supported by the research team. Following screening, children were assessed preintervention, following completion of the intervention and after a 6-month delay. Results: Children in the intervention group showed significantly better performance on measures of oral language and spoken narrative skills than children in the waiting control group immediately after the 30 week intervention and after a 6 month delay. Gains in word-level literacy skills were weaker, though clear improvements were observed on measures of phonological awareness. Importantly, improvements in oral language skills generalized to a standardized measure of reading comprehension at maintenance test. Conclusions: Early intervention for children with oral language difficulties is effective and can successfully support the skills, which underpin reading comprehension. Keywords: Intervention, language, mediation, reading, education.

Introduction

It is well established that learning to read builds on oral language skills. To become literate, children must develop the ability to decode print fluently and the skills to understand what they read; whereas decoding skills depend on phoneme awareness and letter knowledge, broader language skills are required for successful reading comprehension (e.g. Catts, Fey, Zhang, & Tombok, 1999; Muter, Hulme, Snowling, & Stevenson, 2004). Moreover, oral language is a developmental precursor of phonological awareness (Burgess & Lonigan, 1998; Carroll, Snowling, Hulme, & Stevenson, 2003; Storch & Whitehurst, 2002). A corollary of this is that children with language difficulties typically have difficulties with literacy development (Pennington & Bishop, 2009, for review) and are at high risk of educational underachievement.

There is a growing body of evidence-based interventions for children with literacy difficulties (Fletcher, Reid Lyon, Fuchs, & Barnes, 2007; Snowling & Hulme, 2011; for reviews). These studies show that training in the alphabetic principle (letter-sound knowledge and phoneme awareness) combined with text reading are effective in improving word-level decoding difficulties (Hatcher, Hulme, & Snowling, 2004). However, there is variability in response to intervention with poor response being associated with severe phonological impairments (Nelson, Benner, & Gonzalez, 2003; Torgesen et al., 1999; Vellutino et al., 1996) or broader oral language difficulties, including poor vocabulary skills (Al Otaiba & Fuchs, 2006; Duff et al., 2008; Vadasy, Sanders, & Abbott, 2008; Whiteley, Smith, & Conners, 2007). Relatively few studies have targeted reading comprehension, but vocabulary instruction is known to benefit struggling readers (Elleman, Lindo, Morphy, & Compton, 2009) and broader oral language intervention has been shown to be effective for ‘poor comprehenders’ (Clarke, Snowling, True-love, & Hulme, 2010).

Together the findings of these studies provide a rationale for interventions that promote oral language skills to build a secure foundation for literacy, but there is a dearth of evidence concerning their impact (cf. Bernhardt & Major, 2005; Hindson et al., 2005; Nancollis, Lawrie, & Dodd, 2005).

Munro, Lee, and Baker (2008) demonstrated improvements in phonological awareness, vocabu-
lary and oral narrative following a hybrid language intervention targeting vocabulary knowledge and phonological awareness for 5-year-old children, but the lack of a control group limits the conclusions that can be drawn. Pollard-Durodola et al. (2011) randomly assigned 125 children to either receive shared book reading or practice as usual. Significant effects were found in favour of shared book reading, but these were restricted to experimental measures of vocabulary with no transfer to standardized measures, and there was no consideration of the impact of the intervention on literacy development.

Bowyer-Crane et al. (2008) identified children at school entry with poorly developed spoken language skills and provided these children with either an oral language intervention focusing on vocabulary and narrative skills, or a phonology with reading intervention focusing on letter-sound knowledge, phonological awareness and book reading. As expected, the phonology with reading intervention produced improvements in word-level reading skills. In contrast, the oral language intervention produced improvements in vocabulary knowledge and expressive grammar. Similar findings were reported by Bianco et al. (2010) who compared three training programmes; one targeting components of language comprehension (e.g. monitoring, inference making), another targeting language comprehension skills implicitly via storybook reading and a phonological awareness programme. The programme training explicit comprehension skills produced significant improvements in spoken language comprehension, but not in phonological awareness skills, whereas phonological awareness training significantly improved children’s phonological skills, but not their comprehension. However, group allocation was not randomized, and no measure of transfer of gains to literacy development was included.

In summary, although it appears clear that language intervention can be implemented successfully with young children prior to, or at, school entry with positive results, the evidence-base for the effectiveness of oral language intervention starting in preschool is limited and the impact of such interventions on later literacy development remains unclear. The current study evaluated a 30-week oral language programme beginning in Nursery in the term prior to formal school entry (10 weeks) and continuing in the Reception classroom for 20 weeks [The early education system in England is divided into Nursery (ages 3–4) and Primary school (ages 4–11; Reception and Years 1–6). The Early Years Foundation Stage (Nursery and Reception; with the Reception year being pre-Year 1) became part of the National Curriculum in 2002. By law, full-time education is compulsory for all children aged 5–16 and children begin full-time primary education in England by attending a Reception class in the school year they turn 5]. The ‘waiting’ group received no additional teaching during that time.

Children’s performance was assessed at screening (t0) and pretest (t1) and progress was monitored throughout the intervention (t2–t5) and at a 6-month follow-up (t6). Herein, we report data from screening (t0), pretest (t1), immediate post-test (t5) and delayed follow-up (t6). The primary outcome measures were language and spoken narrative skills; the secondary outcomes were phoneme awareness and literacy skills. As a control measure, we also assessed gains in number skills, which were not targeted by the intervention. The timeline for assessments is presented in Figure 1.

Ethical permission for the study was granted by the Research Ethics Committee of the Department of Psychology, University of York. Head teacher consent was given for the intervention to be delivered in schools and informed parental consent was given for each child who participated in the study.

**Methods**

Following screening in 19 preschool settings to identify children with weak oral language skills, we conducted a randomized controlled trial in which children were allocated to either an oral language intervention or a waiting control group. Children allocated to the intervention group received support additional to ‘mainstream’ activities in a 30-week oral language programme beginning in Nursery in the term prior to formal school entry (10 weeks) and continuing in the Reception classroom for 20 weeks [The early education system in England is divided into Nursery (ages 3–4) and Primary school (ages 4–11; Reception and Years 1–6). The Early Years Foundation Stage (Nursery and Reception; with the Reception year being pre-Year 1) became part of the National Curriculum in 2002. By law, full-time education is compulsory for all children aged 5–16 and children begin full-time primary education in England by attending a Reception class in the school year they turn 5]. The ‘waiting’ group received no additional teaching during that time. Children’s performance was assessed at screening (t0) and pretest (t1) and progress was monitored throughout the intervention (t2–t5) and at a 6-month follow-up (t6). Herein, we report data from screening (t0), pretest (t1), immediate post-test (t5) and delayed follow-up (t6). The primary outcome measures were language and spoken narrative skills; the secondary outcomes were phoneme awareness and literacy skills. As a control measure, we also assessed gains in number skills, which were not targeted by the intervention. The timeline for assessments is presented in Figure 1.

Ethical permission for the study was granted by the Research Ethics Committee of the Department of Psychology, University of York. Head teacher consent was given for the intervention to be delivered in schools and informed parental consent was given for each child who participated in the study.

**Participants**

Details of recruitment, selection, allocation and the flow of participants through the study are summarized in

---

In each of the remaining 15 nursery schools, approximately 15 children with the lowest mean verbal composite score were selected as possible participants in the trial. The verbal composite was based on z-scores on screening measures (CELF Preschool IIUK Recalling Sentences, Expressive Vocabulary, Sentence Structure, and Word Structure) were selected to take part in the trial and randomly allocated to either the intervention (N = 90; six from each school) or waiting control groups (N = 90; six from each school). In addition, six children in each school matched on gender and date of birth to a random sample of three children from the intervention and the waiting control groups acted as a representative peer comparison group against which to benchmark the progress of children (N = 82).

**Assessment measures**

The measures tapped skills directly targeted by the intervention and standardized measures of language and literacy. Due to time constraints, not all tests were administered at each testing point.

**Screening tests**

The screening tests (t0) provided an assessment of children’s vocabulary and grammatical development using CELF Preschool IIUK Expressive Vocabulary (requiring picture naming) and Recalling Sentences (requiring sentence repetition), Early Repetition Battery.
word and nonword repetition. Sentence repetition and nonword repetition are established markers of language impairment (Conti-Ramsden, Botting, & Faragher, 2001).

Pre, mid, post and maintenance tests

Language skills.

Grammar: Grammatical skills were assessed using several measures: CELF Preschool II UK Sentence Structure (t1, t3, t5) requires matching one of four pictures to a sentence. The Renfrew Action Picture Test (APT; Renfrew, 2003; t1, t2, t4, t5, t6) measures use of grammar (Grammar Score) and vocabulary (Information Score) for describing pictured actions. CELF Preschool II UK Word Structure (t1, t5) requires the production of inflected forms of verbs and nouns.

Vocabulary: Vocabulary knowledge was measured using the CELF Preschool II UK Expressive Vocabulary test (t0, t4, t5, t6).

Listening comprehension: Listening comprehension was assessed by children listening to two short stories and answering questions about them (t1, t5, t6).

Narrative skills were measured using a story retelling task (Squirrel Story; Carey, Leitao, & Allan, 2006; t1, t2, t4, t5, t6). Narratives were transcribed verbatim and analysed to derive three scores: mean length of utterance in words (MLUw), the number of words used (NW) and the number of different words used (NDW) retelling the story.

Taught vocabulary: The vocabulary taught in the Nursery intervention was assessed using Expressive Picture Naming and Receptive Picture Selection (t1, t2, t4, t5, t6). The vocabulary taught in the Reception intervention was assessed using Picture Naming (t5, t6) and Definitions task (t3, t4, t5, t6).

Phonological awareness.

Onset awareness: Measured using an Alliteration Matching task in which one of two pictures had to be matched to a target picture based on first sound (Carroll & Snowling, 2001). Phoneme awareness: Measured using YARC Sound Isolation in which the initial or final sound has to be identified from spoken non words (t1, t3, t5) and Sound Linkage Segmentation, Blending and Deletion tasks (t6; Hatcher, 2001).

Literacy skills.

Letter-sound knowledge: Knowledge of letter-sounds was assessed at t1, t2, t3, t4, t5, t6.

Reading: The YARC Early Word Reading (Hulme et al., 2009) assessed single word reading (t1, t3, t4, t5, t6). Text reading accuracy and reading comprehension (t6) was assessed using the YARC beginner passage (Snowling et al., 2009).

Spelling was measured by giving children 5 (10) pictures to name and spell (t3, t4, t5, t6). Spelling responses were scored for number of consonants correct.

General cognitive ability. The Block Design sub-scale from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI III UK, Wechsler, 2003) was used as a measure of nonverbal ability (t2).

Intervention programme

Children allocated to the intervention group took part in a 30-week intervention programme delivered by teaching assistants selected by their nursery/school. The first 10 weeks involved three 15-min group sessions (2–4 children per group) per week delivered in preschool.

Once the children entered school, this increased to three 30-min sessions plus two 15-min individual sessions. Children were taught using multisensory techniques within a standard framework; a detailed schedule for the group and individual sessions is presented in Table S1.

The oral language programme aimed to improve children’s vocabulary, develop narrative skills, encourage active listening and build confidence in independent speaking. The programme was a modified version of that shown to be effective by Bowyer-Crane et al. (2008), designed with reference to the UK’s Primary Framework for Literacy and Mathematics (DfES, 2006) and the Statutory Framework for the Early Years Foundation Stage (DCSF, 2008). New vocabulary was selected with reference to themes common in Early Years’ settings (e.g. Growing, Journeys, Ourselves and Time) and included nouns, verbs, adjectives, prepositions, pronouns and question words; taught using multisensory techniques (Beck, McKeown, & Kucan, 2002). Narrative work encouraged expressive language and grammatical competence. Activities revolved around creating and acting out stories, sequencing and story elements.

Listening skills were specifically targeted in the first 20 weeks during the Sound/Listening Game incorporating ideas from Letters and Sounds: Phase 1 (DfES, 2007). This section was extended in the last 10 weeks by activities to promote phoneme awareness (blending and segmenting) and letter-sound knowledge.

Teaching assistants received 2 days training prior to each 10-week block of intervention and attended fortnightly tutorials over the course of the programme. For each 10-week intervention block, a manual described activities and procedures, including materials and resources. To monitor treatment fidelity, teaching assistants attended regular tutorials and the research team observed each teaching assistant delivering intervention and provided feedback on five occasions. In addition, teaching assistants completed records of session plans, children’s progress and attendance for each group and individual session.

Results

Descriptive statistics for all measures at screening (t0), pretest (t1), immediate post-intervention (after 30 weeks, t5) and maintenance test (6 months later, t6) are shown in Table 1 for the intervention and waiting control groups (data for the peer controls is given in Table S2). It can be seen that the intervention and control groups are approximately equated on all measures at t1, as expected given random assignment.

All data analyses were conducted in Stata 12.0 (Stata Corp, College Station, Texas, USA) using ei-
Table 1 Mean raw scores (SD) for intervention and waiting control groups for primary and secondary outcome measures at screening (t0), preintervention (t1), immediately postintervention (t5), and at delayed follow-up (t6; with effect sizes for intervention effects)

<table>
<thead>
<tr>
<th>Primary outcomes</th>
<th>Intervention</th>
<th>Waiting control</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>CELF-EV</td>
<td>t0 (40)</td>
<td>12.60 6.09</td>
<td>12.37 5.97</td>
</tr>
<tr>
<td></td>
<td>t 5 (70)</td>
<td>32.16 10.02</td>
<td>27.84 9.60</td>
</tr>
<tr>
<td></td>
<td>t 6 (70)</td>
<td>36.27 8.54</td>
<td>32.17 9.14</td>
</tr>
<tr>
<td>CELF-SS</td>
<td>t1 (22)</td>
<td>10.15 4.06</td>
<td>10.20 4.45</td>
</tr>
<tr>
<td></td>
<td>t5 (34)</td>
<td>23.45 5.16</td>
<td>22.86 4.50</td>
</tr>
<tr>
<td>APT information</td>
<td>t1 (40)</td>
<td>20.65 6.16</td>
<td>21.06 5.87</td>
</tr>
<tr>
<td></td>
<td>t5</td>
<td>31.40 4.91</td>
<td>29.65 4.88</td>
</tr>
<tr>
<td></td>
<td>t6</td>
<td>31.37 4.73</td>
<td>28.90 5.08</td>
</tr>
<tr>
<td>Listenership comprehension</td>
<td>t1 (16)</td>
<td>3.05 2.43</td>
<td>3.14 2.99</td>
</tr>
<tr>
<td></td>
<td>t5 (16)</td>
<td>6.41 3.34</td>
<td>5.59 3.33</td>
</tr>
<tr>
<td></td>
<td>t6 (14)</td>
<td>7.57 3.00</td>
<td>6.11 2.75</td>
</tr>
<tr>
<td>Narrative MLUw</td>
<td>t1</td>
<td>4.28 1.96</td>
<td>4.74 1.66</td>
</tr>
<tr>
<td></td>
<td>t5</td>
<td>6.81 2.16</td>
<td>6.79 1.78</td>
</tr>
<tr>
<td></td>
<td>t6</td>
<td>7.62 1.95</td>
<td>7.81 2.38</td>
</tr>
<tr>
<td>Narrative NW</td>
<td>t1</td>
<td>50.50 32.77</td>
<td>55.25 34.80</td>
</tr>
<tr>
<td></td>
<td>t5</td>
<td>102.81 47.97</td>
<td>86.58 38.57</td>
</tr>
<tr>
<td></td>
<td>t6</td>
<td>113.15 44.52</td>
<td>101.51 45.10</td>
</tr>
<tr>
<td>Narrative NDW</td>
<td>t1</td>
<td>12.49 7.16</td>
<td>13.27 6.93</td>
</tr>
<tr>
<td></td>
<td>t5</td>
<td>26.23 9.97</td>
<td>23.15 8.85</td>
</tr>
<tr>
<td></td>
<td>t6</td>
<td>27.36 8.86</td>
<td>24.42 9.68</td>
</tr>
<tr>
<td>Secondary outcomes</td>
<td>Alliteration matching</td>
<td>t1 (10)</td>
<td>3.72 2.31</td>
</tr>
<tr>
<td></td>
<td>t5</td>
<td>7.17 2.28</td>
<td>6.59 2.28</td>
</tr>
<tr>
<td>Sound isolation</td>
<td>t1 (12)</td>
<td>0.99 0.36</td>
<td>0.29 0.87</td>
</tr>
<tr>
<td></td>
<td>t5</td>
<td>5.83 3.70</td>
<td>5.46 3.56</td>
</tr>
<tr>
<td></td>
<td>t6</td>
<td>8.42 4.11</td>
<td>7.55 4.32</td>
</tr>
<tr>
<td>Letter knowledge</td>
<td>t1 (17)</td>
<td>1.36 1.70</td>
<td>1.35 2.35</td>
</tr>
<tr>
<td></td>
<td>t5 (17)</td>
<td>13.62 3.68</td>
<td>12.50 3.53</td>
</tr>
<tr>
<td></td>
<td>t6 (32)</td>
<td>27.94 5.59</td>
<td>26.88 5.60</td>
</tr>
<tr>
<td>Early word reading</td>
<td>t1 (30)</td>
<td>0.00 0.00</td>
<td>0.03 0.18</td>
</tr>
<tr>
<td></td>
<td>t5</td>
<td>7.73 6.34</td>
<td>6.68 6.98</td>
</tr>
<tr>
<td></td>
<td>t6</td>
<td>11.94 7.03</td>
<td>11.57 8.73</td>
</tr>
<tr>
<td>Text reading accuracy (errors)</td>
<td>t6</td>
<td>8.57 5.41</td>
<td>8.32 5.84</td>
</tr>
<tr>
<td>Reading comprehension</td>
<td>t6 (8)</td>
<td>4.80 1.58</td>
<td>3.91 1.83</td>
</tr>
<tr>
<td>Spelling</td>
<td>t3 (68)</td>
<td>4.07 5.20</td>
<td>5.42 7.59</td>
</tr>
<tr>
<td></td>
<td>t5 (68)</td>
<td>35.75 18.17</td>
<td>31.78 18.24</td>
</tr>
<tr>
<td></td>
<td>t6 (136)</td>
<td>70.86 30.21</td>
<td>69.94 32.44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>General cognitive ability</th>
<th>Intervention</th>
<th>Waiting control</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>WPPSI block design</td>
<td>.84^d</td>
<td>9.00 2.65</td>
<td>8.91 3.02</td>
</tr>
</tbody>
</table>

[1] Maximum raw scores; RS, recalling sentences; EV, expressive vocabulary; SS, sentence structure; WS, word structure; APT, Action Picture Tests; MLUw, mean length of utterance words; NW, number of words; NDW, number of different words; WPPSI, Wechsler Preschool and Primary Scale of Intelligence.

Reliability: *Cronbach’s alpha; †Interrater reliability; ‡Correlation between parallel test forms; ‡Split-half reliability.

Cohen’s d: ^Difference in progress between groups divided by pooled initial SD; ‡Difference in means divided by pooled SD; t3 Spelling scores given as baseline as not tested at t0 or t1.

other hierarchical linear models or structural equation models (SEM), using Maximum Likelihood Missing Value estimators to allow for missing data and robust (Huber–White) standard errors to allow for the clustering of children within schools. In the SEM models, effect sizes for the intervention effect were calculated from the y-standardized partial regression coefficients for the dummy-coded Group variable. These partial regression coefficients can be interpreted as equivalent to Cohen’s $d$; they express the difference in group means in $z$-score units after allowing for any group differences at baseline (Brown, 2006).

**Effects of intervention on directly taught skills**

Table 2 shows the effects of the intervention on the vocabulary, which was taught in nursery (weeks 1–10) and reception (weeks 11–30) class and on the letter-sounds taught in weeks 21–30. It is clear that there were substantial effects of the intervention on vocabulary taught during weeks 11–30 (effect sizes 0.83–1.18), but much smaller effects on vocabulary taught during weeks 1–10 (effect sizes 0.25–0.27). The intervention also appeared successful in teaching children letter-sounds during weeks 21–30 (effect size 0.41). The $z$ values and significance levels for the hierarchical linear models (Table 2) confirmed that there were no significant effects of the intervention on vocabulary taught during weeks 1–10 (effect sizes 0.25–0.27).

**Generalization to standardized tests**

Our principal interest was to examine the extent to which the intervention produced generalized improvements on tests assessing primary outcomes in the domains of language, and narrative skills and on secondary outcomes, phoneme awareness and literacy skills (see Table 1). To assess these effects we constructed latent variable models in which the four constructs (Language, Narrative, Phoneme Awareness, Literacy) were defined at pretest ($t1$) immediate post-test ($t5$) and maintenance follow-up ($t6$) by multiple measures. The measures used to define each construct at each point in time are shown in the Path diagrams representing these models (Figures 3 and 4). The effects of these analyses were remarkably clear. At immediate and delayed post-test, there were significant effects of the intervention on Language (immediate post-test $d = .80, z = 6.57, p < .001$; maintenance test $d = .83, z = 2.41, p < .001$), Narrative (immediate post-test $d = .39, z = 2.97; p = .003$ maintenance test $d = .30, z = 2.04, p = .041$) and Phoneme Awareness (immediate post-test $d = .49, z = 2.16, p = .031$; maintenance test $d = .49, z = 2.58; p = .01$). The effects on Literacy were not significant (immediate post-test $d = .31, z = 1.81; p = .07$; maintenance test $d = .14, z = .93, p = .354$).

Finally, it is important to note that the intervention group showed higher scores on a reading comprehension test that was first administered at maintenance follow-up. With baseline word-reading skill as the covariate, in a hierarchical linear model with school as a fixed effect, there was a highly reliably effect of the intervention (marginal mean group differ-

| Table 2 Mean raw scores (SD) for intervention and waiting control groups preintervention, immediately postintervention ($t5$) and at delayed follow-up ($t6$) for measures, which were directly taught (with effect sizes) |
|-------------------|-------------------|-------------------|-------------------|-------------------|
|                   | Intervention      | Waiting Control   | Cohen’s $d$       | HLM               |
|                   | $M$               | $SD$              | $M$               | $SD$              | $z$               | $p$               |
| **Intervention vocabulary** |                   |                   |                   |                   |                   |                   |
| Nursery: Expressive naming |                   |                   |                   |                   |                   |                   |
| $t1$ (12)          | .67$^a$          | 1.43              | 1.43              | 1.23              | .27$^1$          | 1.41              | .158              |
| $t5$               | 4.83             | 2.21              | 4.48              | 2.07              |                   |                   |                   |
| Nursery: Receptive |                   |                   |                   |                   |                   |                   |                   |
| $t1$ (12)          | .66$^a$          | 5.64              | 5.64              | 2.16              | .25$^1$          | 1.58              | .114              |
| $t5$               | 9.04             | 1.75              | 8.54              | 2.08              |                   |                   |                   |
| Reception: Definitions |                 |                   |                   |                   |                   |                   |                   |
| $t3$ (84)          | .99$^b$          | 12.90             | 13.47             | 6.98              |                   |                   |                   |
| $t5$               | 29.13            | 9.13              | 21.87             | 7.46              | 1.18$^1$         | 7.17              | <.001             |
| $t6$               | 29.68            | 8.15              | 23.12             | 6.33              | 1.08$^3$         | 6.87              | <.001             |
| Reception: Expressive naming |             |                   |                   |                   |                   |                   |                   |
| $t5$ (24)          | .67$^a$          | 16.94             | 13.72             | 3.79              | .83$^2$          | 6.80$^3$          | <.001             |
| $t6$               | 17.87            | 2.76              | 14.23             | 3.71              | 1.11$^2$         | 9.24$^3$          | <.001             |
| Intervention letter knowledge |           |                   |                   |                   |                   |                   |                   |
| $t4$ (32)          | .91$^a$          | 18.22             | 18.58             | 8.88              |                   |                   |                   |
| $t5$               | 27.80            | 5.71              | 24.18             | 7.35              | .41$^1$          | 4.68              | <.001             |

Reliability: $^a$Cronbach’s alpha; $^b$Interrater reliability.
Cohen’s $d$: $^1$Difference in progress between groups divided by pooled initial $SD$; $^2$Difference in means divided by pooled $SD$; HLM: autoregressor entered as covariate, $^3$t0 CELF EV entered as autoregressor; $t3$ Reception: Definitions scores and $t4$ Intervention letter knowledge scores given as baseline as not tested at t0 or t1.
rence = 0.97, 95% CI 0.40–1.54, \( z = 3.32, p = .001 \). It should also be noted that this effect remained essentially unchanged when reading accuracy on the reading comprehension test at maintenance test was an additional covariate (marginal mean group difference = 0.91, 95% CI 0.42–1.41, \( z = 3.63, p < .001 \)). Thus, the effect of the intervention on reading comprehension is not accounted for by differences in reading accuracy, but appears to be a generalized effect from the improvements in language comprehension skill brought about by the intervention.

To test this hypothesis, we ran a mediation model in which the effect of intervention group on reading comprehension outcome at t6 was the product of improvements in language comprehension abilities at t5 (see Figure 5). In this model, there is complete mediation, with the effect of intervention group on reading comprehension scores being accounted for entirely by the indirect pathway (intervention group \( \rightarrow \) language comprehension t5 \( \rightarrow \) reading comprehension t6). Dropping the direct pathway (intervention group \( \rightarrow \) reading comprehension t6) had no significant effect on the fit of the model (\( \chi^2 \) difference = 0.37, df = 1, NS). The compound path for the indirect effect (intervention group \( \rightarrow \) language comprehension t5 \( \rightarrow \) reading comprehension t6) is highly significant (\( z = 23.68, p < .001 \)).

To check for possible 'Hawthorne' or expectancy effects, arithmetic performance using the BAS Number Concept subtest before and after the intervention was assessed. In a hierarchical linear model with school as a fixed effect, there was no trace of any effect of the intervention on this measure (marginal mean group difference = 0.23, 95% CI 0.00 to 0.45, \( z = 0.66, p = .511 \)).

Table S1 reports data comparing the children in the intervention and waiting control group to the peer comparison group at immediate post-test and maintenance test. These figures allow assessment of the extent to which the intervention had improved language skills to levels comparable to a representative sample of peer controls from the same classrooms. For the measures of language taken at screening and pretest (CELF Recalling Sentences, Word- and Nonword Repetition, CELF Expressive Vocabulary, Renfrew Action Picture Test grammar and information), the differences between the intervention group and the representative peer comparison group are extremely large (\( ds \) range from 1.10 to 1.65). It is clear that where the same measures were reassessed following the intervention, the corre-

Figure 3 Models showing the effects of the intervention on Language (A), Narrative (B), Phonological Awareness (C) and Literacy (D) skills at immediate post-test (t5)
sponding group differences have reduced in size substantially (ds range from .23 to .53) For the remaining tests, for which baseline data are not available, the peer comparison group always show higher scores, but with small-to-medium effect sizes (ds range from .72 to .01). Overall looking at the standard scores and scaled scores shown in Table S1 it is clear that the intervention group are showing essentially normal scores (similar to those of the peer comparison group) on some measures (CELF Expressive Vocabulary, Letter knowledge), whereas on some other measures (CELF Sentence Structure and Early Word Reading) both the intervention and waiting list control group appear to continue to lag behind the peer comparison group. Hierarchical linear models showed that on a number of tests the intervention group did not differ significantly from peer controls following intervention (t6 APT information: marginal mean group difference = -1.28, 95% CI -2.58 to 0.02, z = -1.93, p = .053; t6 APT grammar: marginal mean group difference = -1.06, 95% CI -2.45 to 0.33, z = -1.50, p = .134; t5 letter knowledge: marginal mean group difference = -0.48, 95% CI -1.47 to 0.51, z = -0.95, p = .342; t6 letter knowledge: marginal mean group difference = -0.89, 95% CI -2.39 to 0.61, z = -1.17, p = .243; t6 listening comprehension: marginal mean group difference = -0.79, 95% CI -1.60 to 0.01, z = -1.94, p = .052; t5 narrative NW: marginal mean group difference = -0.46, 95% CI -13.45 to 12.52, z = -0.07, p = .944; t5 narrative NDW: marginal mean group difference = -1.49, 95% CI -4.30 to 1.32, z = -1.04, p = .299; t6 reading comprehension: marginal mean group difference = -0.28, 95% CI -0.88 to 0.32, z = -0.91, p = .364).

**Discussion**

This study evaluated the effectiveness of an intervention designed to improve the oral language and
early literacy skills of children identified in preschool as having poorly developed language. Participating children were randomly assigned to an oral language intervention or a waiting control group. The intervention programme focused primarily on vocabulary, narrative and listening skills, with additional work on letter-sound knowledge and phoneme awareness introduced in the final 10 weeks. The intervention programme was effective in improving oral language and spoken narrative skills immediately following the intervention, and these effects were each maintained 6 months later.

There was evidence that the two preliteracy skills directly taught (letter-sound knowledge and phoneme awareness) benefitted from the intervention. However, these effects did not generalize to measures of word-level literacy skills. This may not be surprising given that these preliteracy skills were only introduced in the last 10 weeks of the programme, and such forms of training are known to be more effective when combined with reading instruction (Hatcher, Hulme, & Ellis, 1994; NRP, 2000).

Notwithstanding the absence of statistically reliable effects of the intervention on word-level literacy skills, a very encouraging outcome was the significant impact of the intervention on reading comprehension measured some 6 months after the intervention had ceased. It is also notable that this effect was statistically independent of any concurrent effects on prose reading accuracy, and the results of a mediation model support the conclusion that this effect is a product of the improved language comprehension skills that were fostered by the intervention (see also Clarke et al., 2010). Furthermore, it is noteworthy that the intervention group was performing at a comparable level to their peers in reading comprehension after receiving the intervention. Arguably, therefore the programme had secured a foundation for reading comprehension in its recipients. More generally, the findings suggest that oral language intervention can make educationally significant improvements to the language skills of children who are ‘at risk’ of failure and it is encouraging that, on the majority of measures, the intervention group had ‘closed the gap’ between themselves and the peer control group after the intervention was completed.

Together the findings of the current trial are very encouraging because the direct effects of the training the children received transferred to standardized measures (cf. Pollard-Durodola et al., 2011). Whereas Elleman et al. (2009) suggested in their meta-analysis of the impact of vocabulary instruction on reading comprehension that standardized measures are often insensitive to vocabulary growth, in this study, its effects were observed in spoken narrative as well as on formal vocabulary measures. However, our data also suggest that the success of vocabulary instruction might depend upon the age at which it is given or the amount of training received: there was no evidence that vocabulary instruction in preschool was effective, although the small effect sizes need to be interpreted cautiously given that exposure to new vocabulary was limited to three sessions per week in the first 10 weeks, compared with daily sessions in the second 20 weeks of the programme when the children were in school. In addition, all teaching in the Nursery part of the programme was in group sessions. It is plausible that the combination of group and individual sessions received during Reception may have been more effective, giving teaching assistants the opportunity to tailor their vocabulary instruction to meet the needs of individual children.

Conclusion
The study has shown that oral language skills can be promoted during the early school years and that this produces effects that generalize to standardized tests of oral language skills and reading comprehension. The study is one of the first to deliver language intervention during the transition from preschool to primary school; a limitation of the design is that it is not possible to evaluate the advantage of starting early (as compared with at school entry) and the optimal timing for and duration of such interventions remain unknown. Nonetheless, the study provides further evidence that school-based oral language interventions can be successfully delivered by trained and supported school staff, a vital step towards developing evidence-based, cost-effective interventions that can be applied effectively in everyday school settings (Savage, Carless, & Erten, 2009). Research is badly needed to extend the current study to evaluate the efficacy of longer term intervention programmes for children with persisting language learning impairments.

Supporting information
Additional supporting information is provided along with the online version of this article.

Table S1. Group and individual session schedule for oral language intervention programme in Nursery and Reception class.

Table S2. Mean raw scores, standard scores and scaled scores (SDs) for peer control, intervention and waiting control groups at screening, pre-intervention, immediately post-intervention (t5), and at delayed follow-up (t6) for key measures (with effect sizes, z values and significance levels from an HLM analysis).

Acknowledgements
The research was funded by a grant from the Nuffield Foundation. We thank Glynnis Smith, Elizabeth Field-
send and Angela Harrington for their help developing the intervention programme and training the teaching assistants, many colleagues for their assistance with data collection, and the teaching assistants, schools, children and parents for their participation.

Key points

- Children who enter school with poorly developed language skills are at high risk of educational failure.
- Data from an RCT show that a 30-week intervention delivered in the last 10 weeks of preschool and continuing for 20 weeks through the transition into primary education can improve oral language, vocabulary and narrative skills.
- Supplementing the language work with training in letter-sound knowledge and phoneme awareness for the final 10 weeks brings about gains in these skills.
- Children who received the intervention showed gains in reading comprehension mediated by gains in oral language (and not by gains in word-level reading skills).

References


Muter, V., Hulme, C., Snowling, M.J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as

Correspondence

Silke Fricke, Department of Human Communication Sciences, University of Sheffield, 31 Claremont Crescent, Sheffield, S10 2TA, UK; Email: s.fricke@sheffield.ac.uk


Accepted for publication: 21 August 2012
Published online: 23 November 2012