Research Report

Is speech and language therapy effective for children with primary speech and language impairment? Report of a randomized control trial

Jan Broomfield† & Barbara Dodd‡
†Speech Therapy 4 Kids, Heighington Village, County Durham, UK
‡City University, London, UK
(Received 3 August 2010; accepted 14 February 2011)

Abstract

Background: A randomized controlled trial was conducted of a whole-service cohort of children referred to the Paediatric Speech and Language Therapy Service of Middlesbrough Primary Care Trust between January 1999 and April 2000. The 730 children involved, aged up to 16 years, were diagnosed with primary speech and/or language impairment in the absence of any learning, physical or sensory difficulty.

Aims: The study considered outcomes of treatment versus no treatment.

Methods & Procedures: Treatment involved clinic-based interventions, provided according to local departmental criteria and care pathways, averaging 5.5 h (range 0–24 h) of contact during the 6-month period immediately following initial assessment. Assessments occurred within 8 weeks of referral and intervention began immediately thereafter.

Outcomes & Results: Treatment was significantly more effective than no treatment, over 6 months, both overall \( (p<0.001) \) and for each of the three diagnostic categories (comprehension, expression and speech, each \( p<0.001 \)).

Conclusions: An average of 6 h of speech and language therapy in a 6-month period can produce significant improvement in performance, and it has been shown to be more effective than no treatment over the same 6-month period for children with primary speech and/or language impairment.

Keywords: children, speech, language impairment, evidence-based practice (EBP), speech and language therapy.

What this paper adds
What is already known on this subject?
There is little prior large group evidence for the effectiveness of intervention for children with primary speech and/or language impairment.

What this study adds
Speech and language therapy intervention is demonstrated as being more effective than no intervention for children with primary speech and/or language impairment.

Introduction

Speech and language impairment may be the most common difficulty presenting in early childhood (Law 1992) with a prevalence of around 6% of children (Bamford et al. 1998). The incidence of referral of this caseload to speech and language therapy (SLT) may be as high as 16.3% of the child population aged up to 16 years (Broomfield and Dodd 2004a). There is still limited evidence, however, concerning the clinical effectiveness of SLT intervention. In addition, the need for early SLT intervention for children who fail to acquire age-appropriate communication skills in the pre-school years has been questioned (Hall 1996). Law et al. (1998) suggested that the benefit is still not proven, and acknowledged that the available studies have been based on individuals and small groups.
Effectiveness of child speech and language therapy

For example, Almost & Rosenbaum (1998) evaluated intervention for 30 pre-school children with severe phonological disorders. They were randomly assigned to two treatment groups. Group 1 received treatment for 4 months followed by 4 months without treatment. Group 2 underwent 4 months without treatment followed by 4 months of treatment. The outcome measures (including phonological error pattern analysis, performance on a standardized assessment, percent consonants correct and mean length of utterance) indicated that group 1 performed better than group 2 on the phonological measures after the first 4 months of the study and in conversational speech intelligibility at the 8-month assessment point. Mean length of utterance scores did not detect a difference between groups at any assessment. The findings of this study reflect those of controlled efficacy studies of intervention for phonological difficulties for smaller group studies and case reports.

A larger-scale study reporting a randomized control trial (RCT) for pre-school children with developmental language disorders was reported by Glogowska et al. (2000). The study involved 159 children, treated by 21 therapists in 16 community clinics. Few differences were found between the treated and untreated groups, although progress made in therapy was greater than that made by children receiving no therapy. The authors concluded that 'this study provided little evidence for the effectiveness of speech–language therapy' (p. 925). One concern is that the description of therapy was limited to 'table-top activities' (Glogowska et al. 2000). The study's outcome measures have also been criticized because the findings included assessment of abilities that had not necessarily been a presenting problem or the focus of therapy (Pring 2004).

Pring’s (2004) critique of the Glogowska et al. (2000) RCT noted that the sample treated was heterogeneous, including children with comprehension difficulties, expressive grammatical difficulties and phonological difficulties, some children having difficulties with more than one aspect of the language. The mean treatment dose was 6 h over 1 year, which is a small amount of therapy according to Conti-Ramsden and Law (2000).

RCTs are considered the gold standard for evaluating clinical interventions across health service provision. Mobley (2000) argued, however, that RCTs were inappropriate in speech–language pathology for several reasons. For example, how outcome is measured determines findings. Some studies measure whether clients have acquired the intervention target in stimuli used in therapy (e.g. Gierut, 2001) as opposed to increased general language competence on standardized assessments; generalizing intervention targets to functional communication; or increased functional communication. When post-intervention performance should be measured is also an important issue. Some skills may show immediate improvement that is not maintained, while others may not show improvement till some weeks after therapy has ceased (Dodd 2007). Another problem identified was that the few meta-analyses available provide little clinically useful information. For example, Law et al.’s (2004) meta-analysis of six studies indicated that SLT intervention resulted in better outcome than no intervention for children with developmental speech disorders. Those six studies, however, used very different types of service delivery (amount of therapy, frequency, length of session) to administer different types of therapy (auditory discrimination, phoneme contrast, articulation of individual speech sounds or not described) to different populations (types of speech disorder, age) using differing outcome measures. Such evidence provides little information concerning case management.

A more recent study (Boyle et al. 2007) reported an RCT of direct versus indirect speech and language therapy for 124 children aged 6–11 years with primary language impairment. The children had non-verbal IQs of more than 75, but performed below –1.25 SD from the mean for their age on the Clinical Evaluation of Language Fundamentals—III UK. They had no sensory or speech difficulties. In addition to assessment, the amount of individual or group therapy given by either a clinician, or trained agents of therapy, was three 30-min sessions per week for 15 weeks. Goals included comprehension, grammar, vocabulary and narrative. The outcomes showed short-term improvements in expressive but not receptive language at the end of therapy. However, the amount of intervention received (22.5 h in addition to assessment time) is considerably more than is typically available in NHS clinics.

There is a need, then, for additional large-scale clinical effectiveness studies to expand the evidence base for the effectiveness of intervention by speech and language therapists (SLTs) for the large population of children with specific speech and/or language impairment. Studies are needed that focus on an entire service rather than on a small number of individual children and SLTs because clinical effectiveness needs to be interpreted in terms of the nature of typical SLT caseloads (Law et al. 1998). Further, there is a need to evaluate particular approaches to therapy (content, amount and service delivery) for specific diagnostic categories (e.g., articulation disorder versus phonological disorder) in well-defined populations using appropriate outcome measures.

The key question addressed here is: does SLT intervention lead to significantly better outcomes for children with specific speech/language impairment than no intervention, over a 6-month period? An additional question concerns whether there are differences in
intervention outcome for comprehension, expressive language and speech impairments. A large randomized-controlled trial of the effectiveness of SLT for all children referred to one paediatric service in Middlesbrough over 15 months was designed:

- to enhance knowledge concerning caseload characteristics, providing a detailed description of all referrals to one children’s SLT service in a given period; and
- to provide evidence about clinical effectiveness of the pathways of care typical of the service’s provision.

The RCT was designed to take place over 12 months for each participant, with assessments at first contact, after 6 months and after 12 months. This paper reports changes made between initial and midpoint assessment, comparing treatment (combining TG1 and TG3) and no treatment (TG2) only. The findings should impact on the future delivery of SLT services to children who have specific speech/language impairment.

The findings of the study at the final, 12-month, point are the subject of a different paper.

**Methods**

**Participants**

The Paediatric section of the Children’s SLT Service of Middlesbrough Primary Care Trust (now Middlesbrough Redcar and Cleveland Community Services) caters for all children who have communication impairment in the absence of any other primary disability such as physical disability, learning difficulty, autism or sensory impairment. Middlesbrough is a large town in North-East England, a locality with high levels of deprivation, and the SLT service caters for Middlesbrough and neighbouring Redcar Cleveland district thereby incorporating both urban inner-city and rural areas.

All referrals received by the SLT department between January 1999 and April 2000 were assessed within 8 weeks of the referral (for further details of the referrals received, see Broomfield and Dodd 2004a). The criterion used to determine specific speech/language impairment was scoring 1 SD below the mean on standardized tests (since this is considered below the statistical normal range) and/or having a profile of clinical symptoms as determined by the assessing SLT in accordance with departmental guidelines. The children invited to participate were assessed using the battery described in Appendix A. Of the 1100 referrals, 740 met the above inclusion criteria and of these 730 consented to take part. Most children were referred by health visitors (50%), followed by teachers (15.2%) and parents (11.5%). Few were referred by general practitioners (4.9%). Boys (70%) were referred more often than girls (30%).

Pre-school children constituted 73.3% of referrals: 6% were less than 2;0; 22% were aged 2;0–2;11; 27% were 3;0–3;11; and 19% were 4;0–4;11 years. Of the 27% of children aged between 5 and 16, most (12% of the whole cohort) were under 6 years and only 2% were older than 11 years. Table 1 shows the age distribution in further detail.

The distribution of socio-economic status was based on the Index of Multiple Deprivation (see http://www.communities.gov.uk/communities/neighborhood/deprivation/), which combines a number of indicators, chosen to cover a range of economic, social and housing issues, into a single deprivation score for each small area in England. The calculation showed that 29% of the population referred was affluent; 29% had low deprivation; 15% had moderate deprivation and 27% had high deprivation. Children were allocated to a primary diagnostic category of comprehension (30.7%), expressive language (25.5%) or speech difficulties (43.8%) by the assessing therapist, based on the assessment findings.

The findings indicated that almost two-thirds of children performed below the normal range (–1 SD on standard scores) on more than one language assessment. The diagnostic group of each participant was determined by the communication skill that was most severely affected. Where two or more areas of functioning were impaired to the same degree, a hierarchy was used to rank functional communication. Comprehension difficulty was ranked above speech difficulty, and both ranked above expressive language difficulty. If a child performed below the normal range on the comprehension measure, they were allocated to the primary diagnostic category of comprehension if no other skills were more impaired than comprehension in terms of standard scores. This decision was taken because difficulties apparent for virtually all these children on expressive language and vocabulary measures might be attributed to their receptive language difficulty. Furthermore, children with comprehension difficulty are at risk of emotional and behavioural difficulties particularly once they start school (Conti-Ramsden et al. 2002). About two-thirds of children with comprehension difficulties also had speech difficulties. Note, however, that means that one third of children with poor comprehension had no speech difficulties, suggesting that the two skills are not necessarily interdependent. About half the children who performed below the normal range on expressive language measures, but had no, or a less severe receptive language difficulty, also had speech difficulties, indicating that expressive
language and speech difficulties can occur independently. Children assigned to the primary diagnostic category of speech difficulties performed within the normal range, or had an impairment of less severity, on receptive language measures. Where expression and speech were impaired to the same degree according to standard scores, children were assigned to the classification of primary speech difficulty, due to the influence that speech disability has on expressive language skills (Dodd 1995).

Children with difficulties pronouncing words may have a difficulty in articulation (ability to learn and make oromotor movements fluently) and/or phonology (ability to acquire the linguistic phoneme contrasts and phonotactics of the phonological system being acquired). Although phonology is a linguistic rather than motor ability, it differs from other expressive language abilities like syntax, morphology and vocabulary in terms of linguistic theory, the ways these abilities are assessed and in the types of intervention implemented for impairment. Phonological impairment can also occur independently of other linguistic impairments (see above). For these reasons, phonology was considered separately from other aspects of language and fell within the speech classification.

**Procedure**

The RCT was designed to take place over 12 months for each participant, with assessments at first contact, after 6 months and after 12 months. This paper reports changes made between initial and midpoint assessment, comparing treatment (combining TG1 and TG3) and no treatment (TG2) only. Table 2 shows the interventions offered during this time, as well as subsequent interventions (in italics).

Intervention was determined by the nature of individual need and based on established departmental care pathways that vary according to age, nature of difficulty and severity. The pathways each identify specific therapy approaches based on both published and practice-based evidence. Therapy programmes implemented were established SLT programmes appropriate for each individual, including Derbyshire Language Scheme (Masidlover and Knowles, 1982), Phonological Contrast Therapy such as Metaphon (Howell and Dean 1991) and Core Vocabulary (Holm et al. 2005). Pre-established departmental criteria determined the specific therapy programme and individual participant targets were set to ensure equity (see Appendix B). The 6-month profile of intervention for each participant consisted of up to four therapy episodes and/or consolidation periods, determined by the study care pathway (see Appendix C). Consequently, participants were offered the type and amount of therapy judged clinically appropriate for their need, with attendance ranging from 0 h (including children offered therapy who did not attend or children given a 6 month review/reassessment period) to 24 h (average = 5.5 h). In most cases, therapy was conducted in small groups (maximum six children) by two speech and language therapists and/or SLT assistants.

### Table 1. Age of participants.

<table>
<thead>
<tr>
<th>Years</th>
<th>Comprehension, n = 224</th>
<th>Expression, n = 186</th>
<th>Speech, n = 320</th>
<th>Total number</th>
<th>Cohort (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; 2</td>
<td>13</td>
<td>25</td>
<td>3</td>
<td>41</td>
<td>5.6</td>
</tr>
<tr>
<td>2 &lt; 3</td>
<td>71</td>
<td>66</td>
<td>25</td>
<td>162</td>
<td>22.2</td>
</tr>
<tr>
<td>3 &lt; 4</td>
<td>63</td>
<td>26</td>
<td>107</td>
<td>196</td>
<td>26.9</td>
</tr>
<tr>
<td>4 &lt; 5</td>
<td>38</td>
<td>13</td>
<td>85</td>
<td>136</td>
<td>18.6</td>
</tr>
<tr>
<td>Pre-school</td>
<td>185</td>
<td>130</td>
<td>220</td>
<td>535</td>
<td>73.3</td>
</tr>
<tr>
<td>5 &lt; 6</td>
<td>14</td>
<td>18</td>
<td>52</td>
<td>84</td>
<td>11.5</td>
</tr>
<tr>
<td>6 &lt; 7</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>40</td>
<td>5.5</td>
</tr>
<tr>
<td>7 &lt; 11</td>
<td>10</td>
<td>24</td>
<td>21</td>
<td>55</td>
<td>7.5</td>
</tr>
<tr>
<td>11+</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>16</td>
<td>2.2</td>
</tr>
<tr>
<td>School age</td>
<td>39</td>
<td>56</td>
<td>100</td>
<td>195</td>
<td>26.7</td>
</tr>
</tbody>
</table>

### Table 2. Study design.

<table>
<thead>
<tr>
<th>Treatment group 1</th>
<th>Treatment group 2</th>
<th>Treatment group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial assessment</td>
<td>Month 0</td>
<td>Month 0</td>
</tr>
<tr>
<td>Phase 1 (0–6 months)</td>
<td>Intervention</td>
<td>No intervention</td>
</tr>
<tr>
<td>Midpoint assessment</td>
<td>Month 6</td>
<td>Month 6</td>
</tr>
<tr>
<td>Phase 2 (6–12 months)</td>
<td>No intervention</td>
<td>Intervention</td>
</tr>
<tr>
<td>Final assessment</td>
<td>Month 12</td>
<td>Month 12</td>
</tr>
</tbody>
</table>
Assessments were conducted by SLTs who were not involved in providing intervention and vice versa, hence blinding was maintained. Initial assessment occurred within eight weeks of referral. Different sets of assessments were used for age bands of children to ensure uniformity (see Appendix A) and were repeated at subsequent assessment points. Assessments were selected based not only on appropriacy of the standardized age range at initial assessment point, but also to accommodate their ability to still be applicable a year on, in order that like with like comparison was possible for each participant; all assessments had a maximum test–retest period of six months. Eligible participants scored at, or below, –1 SD on tests and/or showed a profile of clinical symptoms of speech and/or language difficulties (in accordance with service assessment guidelines); for example, an assessing SLT may find that a child scores within normal limits on assessments but presents with a difficulty such as a lateral /s/—they are thus judged to have a profile of clinical symptoms even though their test scores are within the norm. Children scoring above –1 SD on tests and having no clinical symptoms were classified as having achieved normal functioning. Assessments not only identified the presence of a difficulty and its severity, but also enabled the assessing SLT to identify the main presenting difficulty, whether comprehension, expressive language or speech (articulation and/or phonology). The main assessments used for each age band, based on available standardization, incorporated tests looking at language comprehension, expressive language and speech production. In addition, sampling of oromotor skill, phonological awareness, pragmatic ability and non-verbal functioning took place; again, predetermined tests were used.

Participants were allocated to treatment group, stratified by diagnostic category, off-site through a computer-generated random allocation sequence by the Post Graduate Institute of Health at the University of Teesside. Only the researcher had information linking to consent. Consequently, 730 participants with specific speech/language impairment were recruited, and 710 are reported on at the 6-month point. Figure 1 shows the flow of participants through the study to the midpoint assessment.

The main objective was to ascertain the effect of SLT for children with speech and/or language impairment, comparing children in treatment with those held on a waiting list. In order to account for the effect of maturation, analyses were based on the standardized measurements of each assessment. This involved a range of measures (standard scores on different scales, standard deviations, centiles, quotients), so it was essential to convert all scores for comparison. A z-score is a point on a normal distribution curve that describes how much that point deviates from the mean. z-scores are calculated by subtracting the raw score from the mean score of the sample and dividing that by the sample’s standard deviation. As the nature of intervention targeted the diagnostic category (receptive language, expression or speech), the z-score change for each category is based on the assessment which most reflects such change; so, the z-scores for the receptive language category are based on the receptive language assessment, those for expression are based on the expressive language assessment and those for speech are based on the speech assessment. Consideration of comorbid change is beyond the focus of this paper.

**Results**

Table 3 shows the mean z-scores and their standard deviations for each treatment group at initial and 6-month assessment. Randomization led to the mean z-score start point for the intervention groups (TG1 and TG3) being lower than for the no intervention group, although the range of z-scores for each group was identical (from –3.00 to 0.67) and both groups had significant clinical difficulties. A one-way ANOVA showed that the start-point z-scores were significantly different for the groups receiving intervention than the untreated group ($F_{2,729} = 6.59, p < 0.01$). There may therefore be an element of regression to the mean in TG1 and TG3 which were the ‘in treatment’ groups during the first 6-month period, the focus of this paper; however, the distribution identified above helps to compensate for this. In addition, even the children in TG2 with a mean z-score of –1.48 still had significant improvement opportunity to achieve normal
functioning where \( z = 0 \). Subsequently, the primary measure was change in \( z \)-score between initial assessment and performance after six months of intervention (TG1 and TG3) or no treatment (TG2) on the key assessment administered in relation to the diagnostic category allocated at initial assessment point (see Appendix A). Hence, children with a diagnosis of comprehension difficulty were monitored through comprehension assessment, those with expressive difficulty through the expression assessment and those with speech difficulty through the speech assessment; intervention delivered focussed on this classification, hence measurement focuses on the priority area for each individual.

Figure 2 shows the mean change in \( z \)-score at each assessment point, together with error bars representing 95% confidence intervals, for children who were offered intervention (TG1 + TG3) and those who were not (TG2). The children receiving treatment (\( n = 479 \)) had a mean change in \( z \)-score of 0.67 compared with those receiving no treatment (\( n = 231 \)) who had a change in \( z \)-score of 0.12. The means reflect a statistically significant difference as shown by an adjusted \( t \)-test using \( z \)-score change over time between the treatment and no treatment groups (\( t = 11.58; 660.67 \) d.f.; \( p < 0.001 \)). This significant finding shows that intervention promoted a much greater improvement in speech and language functioning than that occurring through maturation alone. The study has therefore provided evidence that the SLT intervention offered was effective and that, without therapy, children make little progress.

An issue of interest is the percent of children who achieved normal functioning (i.e., performed within one standard deviation of the mean as well as showing no clinical profile of atypical development), both with and without intervention targeting their primary area.
Figure 3. Percentage of participants who achieved normal functioning, improved, maintained age-commensurate progress (static) and fell behind (worse), for treatment and no treatment, between month 0 and month 6 assessment points.

Figure 4. Mean change in z-scores for diagnoses in the treatment and no treatment groups between initial (month 0) and interim (month 6) assessments, with error bars representing the 95% confidence intervals. Comprehension: t = 5.71, 188.22 d.f.; p < 0.001; expression: t = 4.66, 175.64 d.f.; p < 0.001; and speech: t = 10.17, 310.78 d.f.; p < 0.001.

The figure suggests that while treatment promotes greater change than no treatment for each diagnosis, the diagnostic categories do differ. The 147 comprehension cases in treatment had a mean z-score change of 0.81; for the 66 no treatment it was 0.22 (Mann–Whitney U-test, p < 0.001). They made most progress in treatment, but did change without treatment, the difference between the groups being 0.59. The 119 expressive cases in treatment had a mean z-score change of 0.59; for the 63 no treatment it was 0.19 (Mann–Whitney U-test, p < 0.001). They made progress in treatment, but also changed without treatment, the difference between the groups being 0.40. The 211 speech cases in treatment had a mean z-score change of 0.61; for the 101 no treatment it was 0.02 (Mann–Whitney U-test, p < 0.001). They made progress in treatment, but without intervention, hardly changed, and the difference between the groups was 0.59. So, the comprehension group made most overall change in treatment, and treatment made least overall difference to the expressive group.

To investigate the comparative effect of treatment for the three primary diagnostic categories, a series of 2 × 2 contingency Chi-square tests were done. The test comparing the amount of change between treated and untreated groups for comprehension versus expression categories was not significant (Chi^2_1 = 0.228, p = 0.633), indicating that these two diagnostic groups were similar in their response to treatment and no treatment. In contrast, the Chi-square tests indicated that the speech-disorder diagnostic category differed from both the comprehension (Chi^2_1 = 10.452, p < 0.01) and expression diagnostic categories (Chi^2_1 = 12.339, p < 0.001). These findings indicate that untreated children with speech difficulties demonstrate less growth than untreated children in the two language categories.

Figure 5 shows the proportion of children who achieved normal functioning, improved, maintained age-commensurate progress or fell behind, both with and without intervention, within each diagnostic category. The figure shows that regardless of diagnosis, a greater proportion of children in treatment made advanced progress compared with those waiting for intervention. For both comprehension and expression, the profile is similar; however, the profile for children
with speech difficulties appears to differ. To test this observation, the number of children in each of the four categories (normal function, improvement, age commensurate change and poorer performance) was compared in a series of $2 \times 2$ Chi-square tests for each diagnostic category against the other categories, for the treatment group and untreated group separately (table 4). The results showed that there was no significant difference in the proportion of children in the four categories for either the treated or untreated group between comprehension and expression primary diagnostic categories. In contrast, the Chi-square tests indicated that the speech-disorder diagnostic category differed from both the comprehension and expression diagnostic categories, in comparison of both the treatment and non-treatment groups.

### Discussion and conclusion

All children referred to one paediatric speech and language therapy service for whom consent was gained ($n = 730$) participated in an RCT that evaluated specific interventions for well-defined populations of children with speech (phonological/articulatory), expressive language and language comprehension difficulties. The results indicated that an average of fewer than 6 h of intervention was effective. Children with primary speech and/or language disorder who received therapy over 6 months made statistically more positive change in standard scores on widely used speech–language therapy assessments than children held on a waiting list without treatment. Similarly, the results indicated that there was a significant difference in the proportion of children in four outcome categories between the treated and untreated groups. A greater proportion of children in the treated group attained normal functioning or made greater-than-expected improvement for change in age than did children in the untreated group. In contrast, a higher proportion of children in the untreated group performed more poorly than they did at initial assessment or only made change commensurate with age. The findings also indicated that the three primary diagnostic categories (comprehension, expression and speech) showed differences in their response to intervention, and to being left untreated. Children with expressive and receptive language difficulties were more likely to demonstrate growth when no therapy was provided compared with children with speech disorder.

It is worth considering the treatment group separately from the no treatment group. When looking at the treatment group, intervention with the comprehension group made most difference overall (mean $z$-score change of 0.81); this differs from other studies, perhaps due to the young age of the children, the range of severity included and the direct, preset, therapy approach undertaken. Those in the expressive and speech groups made less but still considerable progress (mean $z$-score change of 0.59 and 0.61 respectively). Therefore, overall, the impact of intervention gave an added value of over 0.5 SD for each diagnostic subgroup. When considering the direction of change for the diagnostic subgroups, each had less than 10% making less than age-commensurate progress (comprehension 6.87%, expression 9.2% and speech 1.9%). The language groups had a lower proportion with a static $z$-score (comprehension 12.2%, expression 15.1%) than the speech group (31.8%) and a greater proportion making progress without achieving ‘normal’ (within $–1$ SD of the norm) functioning (comprehension 68.0%, expression 64.7% and speech 46.0%). In contrast, more in the speech group achieved normal functioning

### Table 4. Chi-square test results for the comparison of a number of participants in each diagnostic category who achieved normal function, improved, maintained age-commensurate progress or fell behind for treated and untreated groups.

<table>
<thead>
<tr>
<th></th>
<th>Chi$^2$</th>
<th>d.f.</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment groups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension versus expression</td>
<td>1.931</td>
<td>3</td>
<td>0.587</td>
</tr>
<tr>
<td>Comprehension versus speech</td>
<td>38.341</td>
<td>3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Expression versus speech</td>
<td>38.145</td>
<td>3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Untreated groups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension versus expression</td>
<td>0.195</td>
<td>3</td>
<td>0.978</td>
</tr>
<tr>
<td>Comprehension versus speech</td>
<td>24.739</td>
<td>3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Expression versus speech</td>
<td>25.785</td>
<td>3</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
The no treatment group behaved a little differently. The two language groups, comprehension and expression, made some progress without intervention (mean z-score change of 0.22 and 0.19 respectively). This change was made without any contact or specific recommendations being given to home or daycare settings; it may be that with specific advice, training and target setting, increased gains can be made by this group without direct regular SLT interventions. Again, the majority of the children were of pre-school age, and this may have been a factor in the outcome. In contrast, the speech group made barely any progress overall (mean z-score change 0.02); it should be noted that this means that they kept age-commensurate progress. Again, considering the direction of change, the language groups had around 20% making less than age-commensurate progress (comprehension 22.7%, expression 20.6%), whereas the speech group had 15.87%. The language groups had a lower proportion with a static z-score (comprehension 24.2%, expression 23.8%) than the speech group (64.4%) and a greater proportion making progress without achieving 'normal' (within −1 SD of the norm) functioning (comprehension 47.0%, expression 47.6% and speech 13.8%). In all cases, few children achieved 'normal functioning' (comprehension 6.1%, expression 7.9% and speech 3.0%).

Few previous RCTs into the effectiveness of speech and language therapy have been published and their results have provided contradictory results. The study reported showed that even when all diagnostic categories were combined, intervention resulted in a greater proportion of children making more progress than would be predicted by change in age, in comparison with children who received no treatment. Even though some children in the no treatment group improved while on the waiting list for therapy, the extent of their positive change was statistically significantly less than that made by children who received therapy. These results differ from those of both Boyle et al. (2007) and Law et al. (2003) who reported no positive change for children with language comprehension difficulties. The results also disagree with those of Glogowska et al. (2000) who reported that speech language therapy was of little benefit. Critiques of that study attributed its negative outcome to the heterogeneity of the population, amount of intervention, nature of the outcome measurement and content of therapy (Pring 2004, Dodd 2007). Differences between findings made for large-scale RCTs of paediatric speech and language therapy seem inevitable given the inherent heterogeneous variables. Population-specific factors (e.g., age range, severity, diagnostic categories), intervention-specific factors (e.g., amount and delivery of intervention, skills of the agents of intervention, content of intervention) and research design (e.g., outcome measures, timing of assessments post-intervention) all affect the results found. There seems to be a need for research to address specific questions that investigate the effect of such factors on intervention outcome.

There are a number of possible reasons for the positive difference in findings between this study and others, and indeed these may interact. The factors in the current study that differ from others include:

- The nature of the intervention was tightly prescribed, varying with age and diagnostic profile, with programme specification and resources being supplied, and training given to all SLTs/SLTAs to ensure adherence.
- Clear and specific care pathways were used in order to ensure equity of clinical decision-making.
- Predetermined assessments were used, accounting for age change, such that the same assessment pack was used for individuals at each of the three assessment points.
- Analysis was undertaken using z-scores, based on performance in the area of most significant difficulty (diagnostic classification) at the initial assessment.
- The study population is a clinical population rather than incorporating children within specific parameters of age, severity, a single area of difficulty or other restricting factor.

There is already evidence that diagnostic category affects the outcome of speech language therapy. Law et al. (2003) reported that children with speech disorder responded more positively to intervention than other diagnostic categories, a finding confirmed in this report where double the number of children with speech disorder achieved normal functioning compared with the two language categories. The positive outcome for children with speech difficulties in therapy may reflect the extent of previous research investigating efficacy of intervention (e.g. Weismer, 1981), research on underlying deficits (Stackhouse & Wells 1997) and on differential diagnosis of speech disorder linked to specific types of intervention (Dodd 2005). The results reported here indicated that speech difficulties require direct intervention from SLT services. Children with speech disorder changed little without intervention compared with comprehension and expression categories with the profiles being statistically different for pattern of change. The finding that children in the speech diagnostic category demonstrated little spontaneous change if untreated is surprising. A number of authors have suggested that speech delay, as opposed to disorder,
may resolve spontaneously (e.g. Zhu & Dodd 2000). There are a number of possible explanations for the finding. It might be that 6 months is too short a period for any spontaneous change to make a statistically significant difference. Alternatively, the data might reflect the proportion of children with disordered, rather than delayed, development on the caseload. Also, the differences between this study and prior studies, as listed above, may have a part to play.

Interpretations of outcomes for children with expressive and receptive language impairments are more problematic. Law et al. (1998) reported that young children with language delay made some progress without therapy and that trained agents elicited as much progress as direct therapy from clinicians, a finding replicated by Boyle et al. (2007). One possible interpretation of the findings is that appropriate provision for children with language delay would be indirect input from SLT services: advice, target setting and programme provision. The current study found that children with language impairment, both receptive and expressive, made some progress beyond age-commensurate expectations in the no treatment group, matching Law et al. (1998).

An alternative view is that there is a need to reconsider how children with language impairment are assessed. For example, many currently used assessments fail to discriminate between children with poor performance associated with disadvantaged language learning environments and those who perform poorly despite adequate language-learning exposure (Roy 2010). Alternative assessment methods, such as dynamic assessment, may allow the identification of children whose expressive and receptive language will improve spontaneously (Hasson & Joffe 2007). Further research into the nature of expressive and receptive language difficulties that persist might allow novel and more efficacious interventions to be devised that would target specific underlying difficulties.

Most children who were treated in this study made clinically and statistically significant gains despite being a heterogeneous population in terms of age and diagnostic category and receiving, on average, fewer than 6 h of therapy. Around three-quarters of all children in the treated group made progress that placed them in the normal range for the communication skills targeted (16%), or made progress greater than would be expected for change in chronological age (57%). This positive outcome of intervention might be due to the use of specific interventions for specific, well-defined diagnostic categories using specific pathways of care for children with different intervention needs. Further, more fine-grained analyses of the data need to focus on questions arising from this overview of such a large dataset. For example, it is important to know whether delaying therapy for 6 months affects outcome, and to discover whether the one-quarter of children who failed to make improvements in therapy might have been identified at first assessment.

Acknowledgments

Thanks to the staff of Middlesbrough Primary Care Trust (now Middlesbrough Redcar and Cleveland Community Services) Speech and Language Therapy Service; the data input staff at the Post Graduate Institute of Health, University of Teeside; and to the families and children who took part in the study. The study reported was funded by a National Health Service grant to Dr Jan Broomfield. Ethical approval was granted by the South Tees Nursing and Allied Health Professionals Local Ethics Committee. Funding was by the Northern & Yorkshire NHSE, N046148051, RCTC 66.

Declaration of interest: The authors report no conflicts of interest. The authors are responsible for the content and writing of the paper.

References


DODD, B., ZHU H., CROSBIE, S., HOLM, A. and OZANNE, A., 2002, 
Diagnostic Evaluation of Articulation and Phonology (London: 
Psychological Corporation).

EDWARDS, S., FLETCHER, P., GARMAN, M., HUGHES, A., LETTS, C. 
and SINKA, I., 1997, Reynell Developmental Language Scales, 
3rd edn (Windsor: NFER).

GIERUT, J., 2001, Complexity in phonological treatment: clinical 
factors. Language Speech and Hearing Services in Schools, 32, 
229–241.

GLOGOWSKA, M., Roulstone, S., Enderby, P. and Peters, T., 2000, 
Randomised controlled trial of community based Speech and 
Language Therapy in pre-school children. British Medical 
Journal, 231, 923–926.

GRIFFITHS, R., 1996, Griffiths Mental Development Scales (Bicester: 
Test Agency).

Oxford University Press).

HASSON, N. and JOFFE, V., 2007, The case for dynamic assessment 
in speech and language therapy. Child Language Teaching and 
Therapy, 23, 9–25.

HOLM, A., CROSBIE, S. and DODD, B., 2005, Treating inconsistent 
speech disorders. In B. Dodd (ed.), Differential Diagnosis & 


LAW, J., 1992, The Early Identification of Language Impairment in 

LAW, J., BOYLE, J., HARRIS, F., HARKNESS, A. and NYE, C., 1998, 
Screening for speech and language delay: a systematic review 
of the literature. Health Technology Assessment, 2(9), 1–184.

LAW, J., GARRETT, Z. and NYE, C., 2003, Speech and language 
therapy interventions for children with primary speech and 
language delay or disorder (Cochrane Review). The Cochrane 
Database of Systematic Reviews, 3.

for children with developmental speech and language 
and Hearing Research, 47, 924–943.

MASIDLOVER, M. and KNOWLES, W., 1982, Derbyshire Language 
Scheme (Derby: Derbyshire County Council).


MUTER, V., HULME, C. and SNOWLING, M., 1997, Phonological 
Abilities Test (London: Psychological Corporation).

NEWTON, M. and THOMSON, M., 1976, Good-enough Drawing 
Assessment, Aston Index (Cambridge: LDA).

OZANNE, A., 1992, Normative data for sequenced oral movements 
and movements in context for children aged three to five 
years. Australian Journal of Human Communication Disorders, 
14, 72–83.

PRING, T., 2004, Ask a silly question: two decades of troublesome 
trials. International Journal of Language and Communication 

RENFREW, C., 1995, Word Findings Vocabulary Test (Bicester: 
Winslow).


ROY, P., 2010, Language assessment across social groups: Are our 
tests socially biased? Paper presented at the ESRC Festival of 

SEMEL, E., WIG, E. and Secord, W., 1994, Clinical Evaluation 
of Language Fundamentals (UK) (London: Psychological 
Corporation).

STACKHOUSE, J. and WELLS, B., 1997, Children's Speech and Literacy 

WEINER, F., 1981, Treatment of phonological disability using 
the method of meaningful minimal contrast: two case 
studies. Journal of Speech and Hearing Disorders, 46, 97– 
103.

ZHU, H. and DODD, B., 2000, Development and change in the 
phonology of Putonghua-speaking children with speech 
difficulties. Clinical Linguistics and Phonetics, 14(5), 351– 
368.
Appendix A: Nature of the assessments (Ax)

<table>
<thead>
<tr>
<th>Age range at: I</th>
<th>Initial 0;0 &lt; 1;11</th>
<th>2;0 &lt; 3;5</th>
<th>3;6 &lt; 4;11</th>
<th>5;0 &lt; 6;11</th>
<th>7;0 &lt; 15;11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ax II: 12 months Ax</td>
<td>1;0 &lt; 2;11</td>
<td>3;0 &lt; 4;5</td>
<td>4;6 &lt; 5;11</td>
<td>6;0 &lt; 7;11</td>
<td>8;0 &lt; 16;11</td>
</tr>
</tbody>
</table>

**Main assessments**

<table>
<thead>
<tr>
<th>Language comprehension</th>
<th>REEL</th>
<th>REEL/RDLS</th>
<th>RDSL</th>
<th>RDSL/CEPT</th>
<th>CELF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressive language</td>
<td>REEL</td>
<td>REEL/RDLS</td>
<td>RDSL</td>
<td>RAPT</td>
<td>CELF</td>
</tr>
<tr>
<td>Expressive vocabulary</td>
<td>Word list</td>
<td>Word list</td>
<td>WFVT</td>
<td>WFVT</td>
<td>CELF</td>
</tr>
<tr>
<td>Phonology and consistency of production</td>
<td>Transcript</td>
<td>Transcript</td>
<td>DEAP consistency subtest</td>
<td>DEAP consistency subtest</td>
<td>DEAP consistency subtest</td>
</tr>
</tbody>
</table>

**Supplementary assessments: administered at assessing clinicians discretion**

<table>
<thead>
<tr>
<th>Phonological awareness</th>
<th>n.a.</th>
<th>n.a.</th>
<th>PIPA Sub-tests</th>
<th>PIPA Sub-tests</th>
<th>PIPA Sub-tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatics</td>
<td>Rating scale</td>
<td>Rating scale</td>
<td>Rating scale</td>
<td>Rating scale</td>
<td>Rating scale</td>
</tr>
<tr>
<td>Oromotor skills</td>
<td>Rating scale</td>
<td>Rating scale</td>
<td>Oto-motor</td>
<td>Oto-motor</td>
<td>PAT</td>
</tr>
<tr>
<td>Non-verbal skills</td>
<td>Griffiths</td>
<td>Griffiths</td>
<td>Draw-a-man</td>
<td>Draw-a-man</td>
<td>Draw-a-man</td>
</tr>
</tbody>
</table>

Notes: CELF: Clinical Evaluation of Language Fundamentals (UK) (Semel et al. 1994); Diagnostic Evaluation of Articulation and Phonology consistency subtest (Dodd et al. 2002); Draw-a-Man: Good-enough Drawing Assessment (Aston Index) (Newton and Thomson 1976); Griffith: Griffiths Mental Development Scales (Griffiths 1996); oromotor assessment (Ozanne 1992); PAT: Phonological Abilities Test, speech rate subtest (Muter et al. 1997); PIPA: Preschool and Primary Inventory of Phonological Awareness (Dodd et al. 2000); RAPT: Renfrew Action Picture Test (Renfrew 1997); RDLS: Reynell Developmental Language Scales III (Edwards et al. 1997); REEL: Receptive Expressive Emergent Language Scales (Bzoch and League 1991); WFVT: Word Findings Vocabulary Test (Renfrew 1995); and rating scales for pragmatics and oromotor skills are in-house, developed for use in the study.

Appendix B: Example of departmental clinical criteria for receptive language impairment

<table>
<thead>
<tr>
<th>Age (years; months)</th>
<th>WNL</th>
<th>Action</th>
<th>Mild</th>
<th>Action</th>
<th>Mod</th>
<th>Action</th>
<th>Severe</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>2;0 &lt; 2;8</td>
<td>2 WL</td>
<td>Discharge</td>
<td>1–2 WL</td>
<td>Review, i.e. reassess to check progress made</td>
<td>Early 1 word</td>
<td>Early assessment (EAG) then 1 WL language work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2;9 &lt; 3;2</td>
<td>3 WL</td>
<td>Discharge</td>
<td>1–2 WL</td>
<td>Review</td>
<td>1 WL</td>
<td>1 WL work</td>
<td>&lt; 20 words expressed 1 word level</td>
<td>EAG then early expressive GP 1 WL work</td>
</tr>
<tr>
<td>3;3 &lt; 3;8</td>
<td>3+ WL; early grammar</td>
<td>Discharge</td>
<td>2–3 WL</td>
<td>Review</td>
<td>2 WL</td>
<td>3 WL work</td>
<td>1/2 word level 1–2 WL work</td>
<td></td>
</tr>
<tr>
<td>3;9 &lt; 4;2</td>
<td>4+ WL; grammar</td>
<td>Discharge</td>
<td>3–4 WL</td>
<td>Review</td>
<td>3 WL</td>
<td>4 WL work</td>
<td>1–2 WL work</td>
<td></td>
</tr>
<tr>
<td>4;3 &lt; 4;11 Age appropriate score</td>
<td>Discharge</td>
<td>Score −1 to −1.5 SD</td>
<td>Review</td>
<td>Score around −1.5 SD</td>
<td>High-level language group</td>
<td>−2 SD below the mean</td>
<td>2 WL/3 WL as need plus discuss with the specialist SLT</td>
<td></td>
</tr>
<tr>
<td>5;0 &lt; 6;11 Age appropriate score</td>
<td>Discharge</td>
<td>Score −1 to −1.5 SD</td>
<td>Review</td>
<td>Score around −1.5 SD</td>
<td>High-level language group or IT</td>
<td>−2 SD below the mean</td>
<td>IT plus discuss with the specialist SLT</td>
<td></td>
</tr>
<tr>
<td>7;0 &lt; 16;0 Age appropriate score</td>
<td>Discharge</td>
<td>Score −1 to −1.5 SD</td>
<td>Review</td>
<td>Score around −1.5 SD</td>
<td>High-level language group or IT</td>
<td>−2 SD below the mean</td>
<td>IT plus discuss with the specialist SLT</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Word level (WL) relates to the number of information carrying words a child understands (Derbyshire Language Scheme; Maidloover and Knowles 1982); 1 WL (etc.) refers to therapy focusing at this language level. IT, individual therapy.
**Appendix C: Study care-pathway with examples of treatment programmes**

<table>
<thead>
<tr>
<th>Week</th>
<th>Initial assessment appropriate for age (see Appendix A); duration 90 min for all cohort including TG2 (no treatment group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact for TG1 + 3</td>
<td>Contact according to need and based on clinical criteria (for example, see Appendix B); decision made by the treating SLT at the end of every 6-week episode re: next action</td>
</tr>
<tr>
<td>Weeks 1–6</td>
<td>Six, weekly, 30–60-min sessions of intervention, content based on age and diagnosis, often a small group (moderate/severe on Appendix B) Period for natural maturation of specific skill/area (e.g. expression) (mild on Appendix B)</td>
</tr>
<tr>
<td>Weeks 7–12</td>
<td>Six, weekly, 30–60-min sessions of intervention, content based on age and diagnosis, often a small group Period for natural maturation of specific skill/area (e.g. expression) Review, i.e. reassess area to ensure progress is made</td>
</tr>
<tr>
<td>Weeks 13–18</td>
<td>Six, weekly, 30–60-min sessions of intervention, content based on age and diagnosis, often a small group Period for natural maturation of specific skill/area (e.g. expression) Review</td>
</tr>
<tr>
<td>Weeks 19–24</td>
<td>Six, weekly, 30–60-min sessions of intervention, content based on age and diagnosis, often a small group Period for natural maturation of specific skill/area (e.g. expression) Review</td>
</tr>
</tbody>
</table>

Summary of 6-month period: individuals may have up to four episodes of six, weekly, 30–60-min sessions of therapy targeting their area of need; if they have fewer than four episodes, they may have interim review contacts, where the therapist assessed the area expected to make progress (using a different tool to the study assessment tools). Some participants failed to attend offered appointments, therefore attendance varied from zero to 24 h during weeks 1–24. The mean number of therapy time attended was 5.5 h

Week 25 | Six-month reassessment appropriate for age (see Appendix A); duration up to 90 min for full cohort including TG2. Note: only assessments that fell below –1 SD or showed clinical concern were repeated at this assessment point *(and at the final assessment at 12 months)* |

Examples of 6-week therapy programme (the programme may extend beyond 6 weeks if required, for a variety of reasons, e.g. the child’s absence for 2 weeks, or slower progress than hoped)

<table>
<thead>
<tr>
<th>Week</th>
<th>Two word-level receptive/expressive language</th>
<th>Articulation of /s/</th>
<th>Phonological contrast (Metaphon), e.g. back–front</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Object—place, e.g. Farm topic</td>
<td>Discrimination, context setting</td>
<td>Non-sound level (concepts, representations)</td>
</tr>
<tr>
<td>2</td>
<td>Object—person, e.g. Tea party topic</td>
<td>Isolated sound production</td>
<td>Non-speech sound input (instruments, oral non-speech sound)</td>
</tr>
<tr>
<td>3</td>
<td>Person—possessive, e.g. Dressing topic</td>
<td>Alternate /s/ and vowels; blending /s/ and vowels</td>
<td>Speech sound input</td>
</tr>
<tr>
<td>4</td>
<td>Action—place, e.g. Dolls house topic</td>
<td>CV words; WI /s/ words</td>
<td>Speech sound output; word input (minimal pair)</td>
</tr>
<tr>
<td>5</td>
<td>Action—object, e.g. Bath time topic</td>
<td>VC words; WF /s/ words</td>
<td>Word input and output (minimal pair)</td>
</tr>
<tr>
<td>6</td>
<td>Object—place, e.g. Transport topic</td>
<td>Generalization</td>
<td>Sentence level</td>
</tr>
</tbody>
</table>