The role of language in mathematical development; Evidence from children with Specific Language Impairments.

Chris Donlan  
Dept. of Human Communication Science, University College London

Richard Cowan  
Institute of Education, University of London

Elizabeth J. Newton  
Dept. of Human Communication Science, University College London

Delyth Lloyd  
Dept. of Psychology, University of Melbourne

2950 Words
Abstract
A sample (n=48) of eight year olds with Specific Language Impairments is compared with age-matched (n=55) and language matched controls (n=55) on a range of tasks designed to test the interdependence of language and mathematical development. Performance across tasks varies substantially in the SLI group, showing profound deficits in production of the count word sequence and basic calculation and significant deficits in understanding of the place-value principle in Hindu-Arabic notation. Only in understanding of arithmetic principles does SLI performance approximate that of age-matched-controls, indicating that principled understanding can develop even where number sequence production and other aspects of number processing are severely compromised.

Keywords: Language development; Mathematical development; Specific language impairments (SLI)
The role of language in mathematical development; Evidence from children with Specific Language Impairments.

Evidence from a variety of research areas indicates the involvement of language in mathematical cognition. Spelke & Tsvikin (2001) report language-specific advantages in bilingual adults given training in arithmetic fact retrieval. Exact arithmetic tasks showed benefits only in the language of training, while approximate arithmetic showed equal benefits in both trained and untrained languages. Convergent findings from neuroimaging and ERP studies (Dehaene et al. 1999; El Yagoubi, Lemaire & Besson, 2003), and from brain-damaged patients (Lemer et al. 2003) suggest that the brain-based systems supporting approximate and exact arithmetic may be separable, and that representation of exact number may recruit language-related networks (Dehaene et al 2003; Dehaene et al. 2004).

Recent cross-linguistic studies (Gordon, 2005; Pica et al. 2005), based on languages which lack number words, indicate that exact number representation depends very largely on the availability of a number word sequence, while approximation systems appear to operate independently. Where number word sequences are established, cross-linguistic variation in the structure of the spoken sequence has substantial effects on learning and may influence conceptual understanding (Miura, 1987; Miura & Okamoto, 2003; Miller, Kelly and Zhou, 2004).

The integration of preverbal and verbal systems in the development number processing is currently the focus of much debate. Carey (2004) proposes that linguistic factors play a crucial bootstrapping role in the development of number concepts, through early experience of number-relevant language (Hodent, Bryant & Houde, 2005), and subsequently through integration of the number word sequence with symbolic representations of small sets of items. A contrasting view is offered by Gelman and Butterworth (2005), who propose that numerical cognition is ontogenetically independent, and argue that conceptual understanding does not depend on number word knowledge (Sarnecka & Gelman, 2004).

Landerl, Bevan & Butterworth (2004) examined the role of language in the development of mathematical skills by comparing children with selective deficits in reading or arithmetic, and a dual deficit group, with typically developing children. Performance on a range of basic number processing tasks indicated similar patterns of broad-ranging and substantial impairment in both the arithmetic-only and dual deficit
groups, but not in the reading-only group. These findings suggest that basic number processing deficits underlie arithmetic deficits, and, importantly, that reading deficits do not substantially influence number processing. The close developmental relation between reading and language difficulties (Bishop & Snowling 2004) supports the extended interpretation that language and number are developmentally separable. While acknowledging that phonological aspects of some tasks (number naming and number sequence production) may have affected the performance of their reading deficit group, Landerl et al. (2004) argue that basic number representation (as indicated, for example, by number comparison) is not compromised.

Hanich, Jordan, Kaplan, & Dick (2001) also compared children with selective deficits and typically developing children and have subsequently reported their progress on general measures of achievement and specific numerical tasks (Jordan, Hanich, & Kaplan, 2003; Jordan, Kaplan, & Hanich, 2002). Children with only reading difficulties performed below the level of typically developing children in their understanding of place value, solution of story problems and performance of written computation. In these aspects they resembled children with just arithmetic deficits. Children with both reading and arithmetic difficulties performed worst even when IQ was controlled. Jordan et al. (2003) suggested that language comprehension deficits may inhibit problem-solving, and affect performance on story problems and conceptual understanding of calculation. Mathematical development (Piaget, 1970; Bryant 1995; Baroody, 2003). It is at least plausible to suggest that language, the core medium of teaching, should affect mathematical concepts, though research in the area has focussed more on the relation between procedures (e.g. calculation) and concepts, either as an iterative process.

The development of conceptual understanding is a central issue in mathematical development (Baroody, 2003; Bryant, 1995; Piaget, 1970). It is at least plausible to suggest that language, the core medium of teaching, should affect mathematical concepts, though research in the area has focussed more on the relation between procedures (e.g. calculation) and concepts, either as an iterative process (Rittle-Johnson, Siegler and Alibali, 2001) or as a move from procedural mastery to conceptual understanding (Neches, 1987; Baroody 1995). Neither proposal excludes the possibility that mathematical concepts and procedures are differentially constrained by language, but the issue is complicated by the fact that assessment of conceptual understanding frequently involves self-report or verbal justification.
An important window on the role of language in mathematical development is provided by children with specific language impairments (SLI). These children have significant deficits in expressive and receptive language despite age-appropriate scores on non-verbal ability tests (American Psychiatric Association, 1994). A longitudinal study at ages 5, 7 and 10 (Fazio, 1994, 1996, 1999) found substantial early deficits in production of the number word sequence and cumulative subsequent difficulties in calculation, especially in speeded tasks. At five years the SLI group showed a found a relatively strong grasp of the cardinality principle (whereby the final count word identifies set size), but principled knowledge was not directly tested at follow-up. However, studies using non-verbal response formats have found that seven and eight year olds with SLI exceed the performance levels of language-matched controls in magnitude comparison for single and double-digit numbers (Donlan, Bishop & Hitch, 1998; Donlan & Gourlay, 1999) and matching cardinal values across identity and location change (Donlan, 2003). These findings, based on small samples, suggest that it is at least possible that children with specific language impairments develop conceptual understanding, based on their strengths in non-verbal reasoning, in advance of procedural knowledge compromised by linguistic deficits (Donlan, 1998).

The present study addresses this issue in a large sample of school-age children with SLI. We ask in particular whether language deficits impose a broad ranging obstacle to both procedural and conceptual learning, or whether the non-verbal strengths of children with SLI may support the development of conceptual understanding during the school years. Procedural knowledge is assessed through production of the count word sequence, and performance of basic calculation. Conceptual knowledge is evaluated through understanding of the place-value principle in Hindu-Arabic notation (using multi-digit magnitude comparison) and through understanding of arithmetic principles (using novel stimuli in order to evaluate participants’ grasp of principles independent of their knowledge of specific numerical values, and without the requirement for self-report or justification). We study eight year olds with SLI, compared to a control group individually matched for age, non-verbal ability and school placement (age controls or AC) and to a set of younger controls individually matched with the SLI group for language.
comprehension levels, age-corrected non-verbal ability and school placement (language controls or LC). By selecting controls from the same schools as SLI participants we minimize the effects of environmental variation (Cowan, Donlan, Lloyd & Newton, in press). The design allows us to evaluate the relative contribution of language and non-verbal ability to procedural and conceptual knowledge, and to examine correlational evidence concerning the role played by count sequence knowledge in the development of mathematical skills.

Method

Participants

Participants were 158 children drawn from a pool of 260 attending 23 state schools in locations across Southern England and Wales, excluding major urban centres. All were monolingual English speakers. All children in the SLI group had clinical diagnoses (see Cowan et al., in press). Group measures and inter-group comparisons of language and non-verbal ability are shown in Table 1. All participants completed the experimental tasks described below.

**Counting Aloud.**

There were five different trials: count from one to 41; count backwards from 25; count-on from 25 to 32; count-on from 194 to 210; count-on from 995 to 1010. Number of trials correct, out of five, was recorded.

**Calculation.**

Simple addition and subtraction problems were presented in spoken form. Objects were provided. 16 items were presented in two blocks. The first block comprised 4 addition and 4 subtraction problems (2 + 5, 7 - 5, 2 + 6, 8 - 6, 3 + 6, 9 - 6, 3 + 5, 8 - 5). Testing was discontinued for children who answered all problems incorrectly and for children who became confused or tired. The second block comprised 4 addition and 4 subtraction problems with larger sums and minuends (5 + 7, 12 - 7, 7 + 8, 15 - 8, 8 + 9, 17 - 9, 6 + 7, 13 - 7). Accuracy for each item was recorded.

**Place value principle.**

Understanding of place value was assessed by requiring children to pick the greater of two visually presented multidigit numbers (multi-digit magnitude comparison). The
The role of language

A task consisted of 48 trials that varied in the number of digits in each of the numbers from 2 to 5. In half the pairs the two numbers differed only in one digit, e.g. 45 & 55, 1892 & 1792. A quarter of the pairs contained the same digits in different orders, e.g. 72 & 27, 7431 & 7341, 65984 & 65894. In the remaining pairs the smaller number contain larger digits, e.g. 37& 43, 29996 & 31112, 34343 & 8769. The items were presented on a computer in two blocks. Accuracy was recorded.

Arithmetic Principles.

Children were asked to verify addition and subtraction statements containing unfamiliar numerals. Within a role-play scenario they acted for a Martian maths teacher whose marking of pupils’ homework was interrupted. The test contained 12 trials. Each trial presented a pair of equations. One was already marked as correct (given), the other was for the child to mark (test). In six commuted trials, the addends in the test equation were the commuted version of the given, e.g. given, \( \varphi + \beta = \omega \), test, \( \beta + \varphi = \omega \). In three different trials, one addend in the test equation was the sum from the given equation, e.g. given, \( \sigma + \varphi = \lambda \), test, \( \sigma + \lambda = \varphi \). Three trials involved subtraction. The quantities in the test equation were reversed from the given so the equation could not be correct, e.g. given, \( \beta - \delta = \varphi \), test, \( \delta - \beta = \varphi \). Performance was scored on a categorical basis with ordinal values 0-2, based on responses to trial types. Category 0 was assigned where participants failed to meet criteria for subsequent categories. Category 1 was assigned where participants passed at least 8/9 commuted and different trials. Category 2 was assigned when participants passed at least 8/9 on commuted and different) trials and passed all subtraction trials.

Results

Mean and standard deviations for numerical tasks by group, and test parameters for between group comparisons, are shown in Table 2.

COUNTING ALOUD

All groups showed the same pattern of variation across trials, with rote (0-41) and counting-on (25-32) more successfully accomplished than backward counting, and the higher counting-on trial causing most difficulty. Bonferroni corrected post hoc
comparisons (alpha = 0.05 throughout) showed that AC outperformed both SLI and LC. LC and SLI did not differ.

**Calculation.**

Bonferroni corrected post hoc comparisons showed that AC outperformed both SLI and LC. LC and SLI did not differ. In order to evaluate the mediating effect of Count Aloud scores on the AC/SLI difference in Calculation scores, an ANCOVA was conducted. Count Aloud was a significant predictor of Calculation (F(1, 100) = 37.65, p<.001), and the group difference was abolished (F(1, 100) = 0.420, p = .518)

**Place value principle.**

Bonferroni corrected post hoc comparisons showed that AC outperformed SLI and that SLI outperformed LC. In order to evaluate the mediating effect of Count Aloud scores on the AC/SLI difference in Multi Digit Magnitude Comparison scores, an ANCOVA was conducted. Count Aloud was a significant predictor of Multi Digit Magnitude Comparison (F(1, 100) = 35.36, p<.001), but the group difference was abolished (F(1, 100) = 0.588, p = .445). A further comparison of AC vs. SLI performance on double-digit stimuli only confirmed the group difference (F(1,101) = 34.14, p<.001). ANCOVA showed that Count Aloud was a significant predictor of Double Digit Magnitude Comparison (F(1, 100) = 37.65, p<.001), but the group difference was abolished (F(1, 100) = 0.420, p = .518).

**Arithmetic Principles**

Frequencies of response category by group are shown in Table 3.

The response range was narrow. Only three participants (all from the AC group) showed full understanding of addition and subtraction principles. 53 out of 55 children in the LC group failed the task altogether. One-way ANOVA examined the effect of Group on knowledge of Arithmetic Principles. The effect was significant (F(2, 155) = 14.84, p<.001, Eta Sq = 0.161). Bonferroni corrected post hoc comparisons showed that both AC and SLI outperformed LC, but that AC and SLI did not differ. In order to evaluate the mediating effect of Count Aloud scores on the AC/SLI group contrast, an ANCOVA was conducted. Count Aloud scores were unrelated to knowledge of Arithmetic Principles (F(1, 100) = 1.39, p = .24), and the effect of Group was not significant (F(1, 100) = 0.002, p = .964).
Discussion

The finding that children with SLI have severe deficits in both counting and calculation is not new (Fazio 1994, 1996, 1999). The extent of the counting deficits may be surprising; error analyses revealed that forty percent of the SLI group, but only four per cent of language controls, failed to count to twenty. In calculation, though presentation of problems was in spoken form only, performance levels are entirely consistent with those found for visually presented problems (Fazio (1996, 1999; Cowan et al., in press). Statistically, the calculation deficit in our SLI group is fully explained by performance in verbal counting, but a cautious interpretation is required here. Causal linkage cannot be inferred. Nor is there a literal correspondence between the specific range of numbers processed in each task. No child in the study failed to count to ten, but the SLI group performed significantly more poorly than age controls in every calculation trial, even where sums or minuends fell below ten.

In understanding of the place-value principle (multi-digit magnitude comparison) SLI performance was in deficit compared to age controls, but significantly exceeded the level of language matched controls. This lends some support to the proposal of Donlan and Gourlay (1999) that understanding of the place-value principle may be language-independent. We note the correlational evidence that the SLI deficit can be fully explained by verbal count performance.

Perhaps surprisingly, our stringent test of arithmetic principles (a more abstract test of principled knowledge than those used by Jordan et al., 2003) shows no clear deficit in the SLI group, relative to age controls. Many children with SLI are as capable as typically developing peers of grasping the logical principles underlying simple arithmetic, despite substantial procedural deficits. Of the 19 individuals with SLI who showed understanding of arithmetic principles, 10 failed to count correctly to 41.

These findings challenge previous accounts of the development of arithmetic knowledge (Piaget 1964; Baroody, 1995; Rittle-Johnson et al., 2001) by suggesting that conceptual understanding may be achieved despite severe procedural deficits. Recent work by Canobi (in press) indicates a possible explanation. Canobi classified a subset of her sample of seven to nine year olds as ‘symbolic thinkers’ capable of abstract reasoning about addition and subtraction problems, and more likely to demonstrate conceptual understanding in a symbolic than a concrete context. Our test
of arithmetic principles is maximally symbolic, since the ‘numerals’ employed have no specific values. This may facilitate the detection of conceptual relations for children whose ability to manipulate actual numerals may be impaired. An account of this sort is compatible with the proposal of Huttenlocher, Jordan and Cohen Levine (1992) that non-verbal mental models, emerging between the ages of two and three, may provide the basis for young children’s arithmetical reasoning. Our findings suggest that, for some individuals, ‘non-verbal calculation’ (Huttenlocher et al., 1992, p.295) may continue to develop during the school years. On the other hand, in contrast to the findings of Landerl et al. (2004), we find that language impairments present substantial obstacles to the development of conventional arithmetic procedures. It is important to emphasise that the current sample differs from that of Landerl et al. (2004), and from that of Jordan et al. (2003) insofar as the language-impaired group is selected on the basis of language deficits rather than reading deficits, and shows particularly poor performance in both phonological and grammatical processing. Further work is needed to explore the possible underlying relations between linguistic and numerical systems which these findings suggest.

Conclusion

Specific language impairments in childhood inhibit acquisition of the spoken number sequence, development of calculation skills and, to a lesser extent, acquisition of the place-value principle in Hindu-Arabic notation. Nonetheless, acquisition of the logical principles of simple arithmetic may be unaffected. The linguistic constraints which regulate children’s developing ability to produce the spoken number sequence may affect the development of conventional calculation skills and understanding of number notation. However, the development of knowledge of arithmetic principles may be supported by a separable system.

Acknowledgement: The study reported here was completely funded by the Nuffield Foundation, London, UK.
Reference List


Table 1 Characteristics (mean, s.d.) of the Language Control (LC), Specific Language Impairment (SLI), and Age Control (AC) Groups for Nonverbal Ability and Language with Group Comparison Test Parameters.

<table>
<thead>
<tr>
<th>Measures</th>
<th>LC</th>
<th>SLI</th>
<th>AC</th>
<th>MSe (df = 2, 155)</th>
<th>F Ratio</th>
<th>p-value</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronological age (years)</td>
<td>6.0 (0.4)</td>
<td>8.3 (0.4)</td>
<td>8.2 (0.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven(^1), IQ</td>
<td>106.6 (10.9)</td>
<td>103.0 (12.3)</td>
<td>104.6 (11.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven(^1), Raw Score</td>
<td>18.4 (4.0)</td>
<td>24.2 (4.7)</td>
<td>24.8 (4.5)</td>
<td>19.2</td>
<td>35.6</td>
<td>&lt;.001</td>
<td>.31</td>
</tr>
<tr>
<td>TROG(^2), Standard Score</td>
<td>94.6 (7.2)</td>
<td>80.6 (6.4)</td>
<td>100.0 (11.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TROG(^2), Raw Score</td>
<td>11.7 (1.7)</td>
<td>11.5 (1.7)</td>
<td>15.9 (1.78)</td>
<td>2.9</td>
<td>115.6</td>
<td>&lt;.001</td>
<td>.60</td>
</tr>
<tr>
<td>Non Word Repetition(^3)</td>
<td>22.5 (5.8)</td>
<td>11.2 (5.7)</td>
<td>27.0 (4.5)</td>
<td>28.9</td>
<td>115.9</td>
<td>&lt;.001</td>
<td>.60</td>
</tr>
<tr>
<td>Past Tense Production(^4)</td>
<td>10.7 (2.8)</td>
<td>5.5 (4.0)</td>
<td>15.8 (2.6)</td>
<td>10.18</td>
<td>132.2</td>
<td>&lt;.001</td>
<td>.63</td>
</tr>
</tbody>
</table>

NB Power=1.0 for all comparisons

\(^1\) Raven’s Coloured Progressive Matrices, Raven (1998)
\(^2\) Test for the Reception of Grammar, Bishop (1983)
\(^3\) The Children’s Test of Nonword Repetition, Gathercole and Baddeley (1996)
\(^4\) Task adapted from Marchman, Wulfeck and Ellis-Weismer (1999)
Table 2. Performance (mean, s.d.) of the Language Control (LC), Specific Language Impairment (SLI), and Age Control (AC) Groups on Numerical Tasks, with Group Comparison Test Parameters.

<table>
<thead>
<tr>
<th>Measures</th>
<th>LC</th>
<th>SLI</th>
<th>AC</th>
<th>MSe (df = 2, 155)</th>
<th>F Ratio</th>
<th>p-value</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>55</td>
<td>48</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Count Aloud (max=5)</td>
<td>1.7 (1.1)</td>
<td>1.7 (1.4)</td>
<td>4.1 (1.0)</td>
<td>1.4</td>
<td>70.3</td>
<td>&lt;.001</td>
<td>.47</td>
</tr>
<tr>
<td>Calculation (max=16)</td>
<td>7.2 (4.9)</td>
<td>8.9 (4.5)</td>
<td>13.4 (3.4)</td>
<td>18.3</td>
<td>31.3</td>
<td>&lt;.001</td>
<td>.29</td>
</tr>
<tr>
<td>Multidigit Magnitude</td>
<td>31.8 (6.5)</td>
<td>35.5 (6.6)</td>
<td>42.1 (4.7)</td>
<td>35.5</td>
<td>41.3</td>
<td>&lt;.001</td>
<td>.35</td>
</tr>
</tbody>
</table>

NB Power=1.0 for all comparisons
Table 3. Frequency of response category by group for the Arithmetic Principles task

<table>
<thead>
<tr>
<th>Measures</th>
<th>LC</th>
<th>SLI</th>
<th>AC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 0 (fail)</td>
<td>52</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>Category 1 (pass commuted and different trials)</td>
<td>3</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>Category 2 (pass commuted, different and subtraction trials)</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

$n$ = 55, 48, 55