Risk markers for SLI: a study of young language-learning children

Gina Conti-Ramsden and Anne Hesketh
Human Communication and Deafness, University of Manchester, Manchester, UK

(Received 30 July 2002; accepted 24 January 2003)

Abstract

Background: Adequate assessment of ‘normal variation’ versus ‘abnormal status’ is particularly difficult for clinicians working with young children who are under 5 years of age and who present with slow language development. There is therefore clinical motivation to identify possible key difficulties (or ‘risk markers’) that may distinguish children who are likely to have specific language impairment (SLI) from the variation observed in younger, normally developing children.

Aims: The issue of ‘risk markers’ for SLI was explored. It is well known that the issue of markers is controversial. The view presented here is that a risk marker represents a symptom with no assumption about whether the symptom reflects a single cause or that this symptom alone identifies the disorder.

Methods and Procedures: The performance of 32 children with SLI was compared with that of 32 younger normal language-learning children on four potential risk marker tasks: non-word repetition, digit recall, past tense provision and plural marking.

Outcomes and Results: The findings suggest that processing markers, particularly non-word repetition, have the potential for indicating SLI risk. In particular, children who fall in the bottom quarter of the normal distribution in non-word repetition (performance below the 25th centile) appear to be at risk of SLI.

Conclusions: Although these results need to be interpreted cautiously, the evidence is thought to warrant the use of non-word repetition tasks in clinical practice.

Keywords: specific language impairment (SLI), risk markers, non-word repetition.

Introduction

Clinical experience suggests that when dealing with young language-learning children, say 5 years of age or younger, the assessment of possible language difficulties involves a challenging and complex set of issues (for a recent review, see Leonard...
Clinicians establish (usually in collaboration with other professionals) the hearing status of children as well as their general non-verbal abilities with the view to discounting hearing impairments or general learning difficulties as possible causal factors. There is no doubt that as far as language abilities are concerned, clinicians first take into consideration the child’s age and make appropriate comparisons for the child’s language abilities with respect to the ‘norm’ using standardized or non-standardized tools, systematic observations, language samples (amongst other instruments) used singly or in combination in the evaluation process. In addition, clinicians often make judgments about children’s profiles and behaviours and, in doing so, draw on their knowledge of normal language (NL) development to ascertain whether the children in question are behaving much like younger NL-learning children or whether they are different in any way. This latter process of evaluation involves the clinician making comparisons between possible affected children and younger normally developing children.

Perhaps due to the above, much research in specific language impairment (SLI) has involved comparisons with two groups of children, both chronological age matches as well as language stage matches (using a variety of measures to match the language abilities such as mean length of utterance (MLU), expressive language, receptive vocabulary, etc.). Comparisons between children with SLI and younger language-learning children are complex, particularly when they involve language or language-related measures, as these are usually continuous variables that change with development. Nonetheless, a general picture has emerged that children with SLI are very similar to normal, younger language-learning children.

This idea is worth exploring further and at least one reason comes from what we know about NL development itself in children 5 years of age and under. It is known that there are large individual differences in the rate of language development and in learning style in normally developing children (Bates et al. 1995). This variation is thought to be stable and substantial, making the adequate assessment of ‘normal variation’ versus ‘abnormal status’ particularly difficult for clinicians working with young children. There is therefore an added clinical motivation to identify possible key difficulties (or ‘risk markers’) in SLI that may distinguish these children from the variation observed in younger, normally developing children.

What do we mean by a ‘risk marker’? There are at least two possible interpretations of what a marker for a disorder should be. A strong interpretation suggests that a marker represents not only a clear symptom of a problem, but also a particular cause. A less strong interpretation takes a marker to represent the clearest symptom, with no assumption about whether the symptom reflects a single cause. Research in the area of SLI is controversial with regard to the issue of markers. Nonetheless, investigators have used both a strong interpretation of markers (e.g. Rice et al. 1995, 1998, Rice and Wexler 1996) as well as less strong interpretations (Bishop et al. 1996, Dollaghan and Campbell 1998).

An even milder interpretation is what we want to argue for here. This interpretation suggests that a risk marker represents a symptom, with no assumption about whether the symptom reflects a single cause nor that this symptom alone identifies the disorder. On the contrary, it is assumed that the risk marker is more likely to be used in combination, to complement information available. Concretely, we are thinking of the common clinical situation where a child is already presenting with a language problem and the clinician is attempting to make a decision about whether the child is likely to be part of the normal continuum and variation seen in young language-
learning children or whether the child is at risk for SLI. Could there be ‘risk markers’ that may help clinicians identify children who are more likely to be SLI?

The search for strong clinical markers for SLI has traditionally been divided into two main approaches: research based on processing considerations versus research based on linguistic frameworks. In the processing approach, the role of memory in language performance, particularly performance on short-term memory tasks such as non-word repetition, has been found to be a good indicator of SLI (Gathercole and Baddeley 1990, Bishop et al. 1996, Dollaghan and Campbell 1998, Ellis Weismer et al. 2000). In the linguistic approach already mentioned, the ability to mark syntactic tense in English has been proposed as a strong clinical marker as it appears to be a particular area of difficulty for children with SLI throughout the primary school years (Rice and Wexler 1996, Rice et al. 1995, 1998, Marchman et al. 1999). This division in research approaches has resulted in a dearth of studies that examined more than one type of marker in parallel, i.e. studies that analyse the sensitivity and specificity of each marker and the possible contributions of combinations of markers to the identification of SLI (but see Conti-Ramsden et al. 2001 for an exception with older children, and Conti-Ramsden 2002 for comparisons between 5-year-old children with SLI and chronological age controls). Thus, a secondary aim of the present study was to examine a number of markers and the possible contribution of combination of markers to the identification of SLI risk.

Non-word repetition and linguistic tense having been identified as two likely markers for young children with SLI, the literature was examined for other possible candidate markers for the disorder. Many studies suggest that grammatical morphemes are particularly vulnerable in children with SLI. Although most of the reliable differences have been found for grammatical morphemes for tense and agreement (verb morphology), there is some evidence that noun morphology may also be problematic. For example, Bedore and Leonard (1998) carried out a discriminant function analysis in which they included a noun morpheme composite (as well as a verb morpheme composite and MLU in morphemes). They found that in a small group of 12 children aged 4;0–5;8 years, both the noun and the verb composites (separately as well as together) were highly accurate in classifying children with and without SLI. In the same area, Oetting and Rice (1993) found differences between 5-year old children with SLI and age-matched peers on the ability to pluralize nouns (both frequently and infrequently pluralized nouns). Similarly, Crago and Gopnik (1994) and Gopnik and Crago (1991) examined the use of grammatical morphology by 20 members of a three-generation family, 13 of whom exhibited language problems. Deficits in the use of noun plural -s (among other morphemes) were observed in the affected members of this family. However, such results have not always been consistent, even for young children with SLI (e.g. Rice and Oetting 1993, Rice and Wexler 1996). It was clear that further examination of noun pluralization in SLI would be of interest and so plural marking was included as a possible risk marker of SLI in the present study.

Finally, we examined other possible processing tasks, particularly those involving short-term memory. Ellis Weismer et al. (2000) suggest that tapping working memory via tasks such as the sentence span task (Daneman and Carpenter 1980) or the competing language processing task (Goulin and Campbell 1994) has potential for differentiating 7-year-old children with SLI from NL learners. However, such tasks were likely to be too complex for the younger children of interest to this study (2–5 years) and so digit recall was chosen as the second processing marker. Although
not previously investigated as a potential marker, this task has typically been used as a measure of short-term memory (Baddeley et al. 1998) and is very easy to administer to young children. It assesses the maximum sequence of digits a child can remember correctly.

Thus, four potential risk markers were examined in the present study. Two processing markers: non-word repetition (Gathercole and Baddeley 1990) and recall of digits (subtest of the British Ability Scales; Elliot 1983) and two linguistic markers: past tense marking (Marchman et al. 1999) and plural marking (Oetting 1992, Oetting and Rice 1993). In this paper, we specifically focused on comparisons between children with SLI and younger NL-learning children.

Methods

Participants

Thirty-two children with SLI and 32 younger NL control children participated in the study. All children were monolingual English speakers and were from the North West of England and from London. Children with SLI were recruited through speech and language therapists or specialist teachers in language units. NL control children were recruited from the same area as the children with SLI. Children were not included if they presented with any severe behavioural, hearing or speech problems (noticeable articulatory, dyspraxic or phonological problems) as assessed by an interview with their teacher and/or speech and language therapist. All children with SLI were receiving specialist teaching for SLI, either within a language unit or in mainstream education, with learning support teachers.

The group of children with SLI included seven girls and 25 boys, ranging in age from 52 to 70 months, with a mean age of 61 months (5;00 years). The language ages and standard language scores of the children with SLI were established via the CELF-Preschool (Wiig et al. 1992; American standardization). The language age in this group of children ranged from 30 to 46 months, with a mean of 36 months (3;00 years). Standard scores for the children with SLI ranged between 50 and 82 (mean = 69.9, SD = 7.41). In addition, all children with SLI had an IQ >80, as measured by the WPPSI-R (Wechsler 1992) performance IQ subtests (mean = 95.84, SD = 11.11).

Children with NL were recruited through nurseries and primary schools. The group of children with NL development comprised 12 girls and 20 boys. Children with NL ranged in age from 28 to 43 months, with a mean chronological age of 34 months (2;10 years). The language age in this group of children ranged from 29 to 45 months, with a mean language age of 34 months (2;10 years). The language ages of the children with NL was determined using the Reynell Developmental Language Scales — III (Edwards et al. 1997; British standardization). The decision to use different tests was taken given the younger age range of the NL group (28–43 months) and the fact that the CELF-Preschool was designed for children 3;0 years or older. At the beginning of the study we focused on recruiting the children with SLI and used the CELF-Preschool. On the second phase, when we recruited the younger NL-learning children, it became clearer that we would have to engage with children younger than 3;0 years in order to obtain a rough language age match group. Thus, the decision to use different tests involved considerations of age-appropriateness of the tests as well as practical considerations. The language ages of the children with NL were within 1 month of their chronological ages, ensuring
they were age appropriate with regards to language skill. All children with NL had an IQ > 80, as measured by the WPPSI-R (Wechsler 1992) performance IQ subtests (mean = 100.59, SD = 11.73). The WPPSI-R was designed for children from the age of 2;11 upwards. However, it was decided to use the WPPSI-R with all the children with NL, even if they just fell below the lower age limit. Children falling below the age limit were scored as if they were 3 years of age which provided a conservative estimate of their non-verbal abilities.

No attempt was made to match on gender as priority was given to the variables of non-verbal IQ and language age. Having said this, the distribution of gender was not significantly different for the two groups of children ($\chi^2 = 1.87, p > 0.10$). Furthermore, there were no significant differences between the groups on performance IQ ($F(1,62) = 2.77, p > 0.10$).

In addition, the children with SLI and the younger NL controls had similar language ages. It needs to be pointed out, however, that there are known difficulties in the use of age equivalent scores (relationships between age and test means and variances are seldom constant) and these difficulties may have been exacerbated by the use of two different tests of language in this study (for a review of the problems of using age equivalent scores, see Bishop 1997). Nonetheless, the aims of the present study required only a rough estimate of language development for which language age equivalent was thought to be robust enough. The important point was that the younger NL control group was on average at least 2 years younger than the SLI group.

Procedure

Following receipt of informed written consent from parents, children were assessed individually in a quiet room within their school or nursery. Testing on risk marker tasks was completed as part of a wider battery of language and cognition tests, during several visits, at the child’s own pace and with normal school breaks.

Testing was carried out by a single researcher, thus no measures of reliability of the coding/transcription of responses were available. Note that the researchers involved in data collection were highly experienced (had been working with young children and children with SLI for at least 2 years) and had received extensive training on the tasks used.

Risk marker tasks

Past Tense Task (PTT)

A test of past tense, designed to assess grammatical usage of verbs in the past tense, was administered. The test was developed by Marchman for younger children (3–6 year olds), was a modified, shortened version of the PTT developed by Marchman et al. (1999), and comprised 30 regular and irregular verbs. The child was shown a black-and-white picture depicting everyday activities while the experimenter read a sentence, omitting the target verb. For example, children were shown a picture of a girl kissing her father, while simultaneously hearing: ‘The girl is kissing her father goodnight. She kisses him every night before bed. Last night she _____ (kissed) him’. The experimenter attempted to elicit the past tense of the verb ‘kiss’ from the child, using a flat but lengthened contour on the last word before the pause to suggest that the sentence was incomplete. Irregular verbs were also used in the task. For example, children were shown a picture of a girl riding
her bike while the experimenter read aloud: 'This little girl is riding her bike to school. She rides her bike to school every day. Yesterday, she ___ (rode) her bike to school'. One point was awarded for each verb that was correctly marked for tense. This point was also awarded in cases of child over-regularizations (e.g. 'rided') or changes in the lexical item (e.g. 'touched him' when the target was 'kissed'). Scores were calculated from the total number possible, and a proportion correct derived from that score. No standardized information is available for this task.

**Noun plural task**

A test of noun plurals was administered, following a task originally designed by Oetting (1992) and Oetting and Rice (1993). The task used here was the modified, shortened version developed by Marchman (personal communication 2001). The test comprised 31 nouns and included regular and irregular nouns. For the regular plural nouns, items were divided into /-s/, /-z/, and /-lz/ categories. Two frequency bins were created based on the LEX database for American English, the item frequencies were taken from the norming study of the MacArthur CDI (Fenson et al. 1991). Children were shown black-and-white pictures of nouns (the original task by Oetting used picture books). Each noun was shown singularly (e.g. one flower) on one half of the page, while on the other half of the page, the plural of the noun was represented (e.g. two flowers). For example, children were shown a picture of one flower while the experimenter said aloud: 'Look — here is one flower. However, over here are two ___ (flowers)'. One point was awarded for each noun that was correctly marked. Scores were calculated from the total number possible, and a proportion correct was derived. Note that on the original task by Oetting, the plural pictures involved six referents instead of two, and non-numeral quantifiers (e.g. 'you have many ___'), whereas we used numerals. No standardized information is available for this task.

**Non-word repetition**

The Children's Test of Non-word Repetition (CNRep: Gathercole and Baddeley 1996), a test of phonological short-term memory comprising 40 non-words of two to five syllables, was administered live by the experimenter. The experimenter said a non-word aloud, obscuring the view of his/her lips while saying the non-word using a stuffed toy. Non-words were uttered only once. Following this, the child was asked to repeat the stimuli as accurately as possible. As recommended by Gathercole and Baddeley, this test was scored 'on line' with each item judged as correctly or incorrectly repeated.

As mentioned above, all children with noticeable speech problems were excluded from the study. In addition, an informal conversation was carried out with the children in order to assess their phonology and note any segmental substitutions that were based on normal phonological processes (e.g. fronting 't' for 'k'). Children who at this stage were difficult to understand were excluded from the study. Thus, segmental substitutions based on normal phonological processes were not counted as incorrect for any of the children participating in the study. Prosodic processes such as omission of prestressed syllables were counted as errors. Each item correctly repeated was awarded one point, resulting in a total score of a number correct out of 40. This test has been standardized for children aged 4–8 years. A proportion of the children participating in this study were under the age of 4 years, thus we
used regression analysis to extrapolate the centile scores for 2 and 3 year olds for the 10th, 25th, 50th, 75th and 90th centiles.

**Digit recall**

This aspect of auditory verbal short-term memory was assessed via the Recall of Digits subtest from the British Abilities Scale (BAS) (Elliot 1983). Children were presented with a sequence of numbers, starting with a list length of two (e.g. 3, 5). They were asked to repeat the sequence of numbers spoken aloud by the experimenter in the same order as presented. A block of five number sequences with a length of two was read out (e.g. 4, 4; 2, 3; 5, 4; 9, 2; 7, 5) first. The following block of five number sequences increased in length by one (e.g. 8, 6; 2, 4, 2, …). The maximum length of a number sequence was nine numbers. The experimenter continued until the child failed to list correctly five consecutive number sequences. This test was standardized for children aged 2;6–17;11 years.

**Results**

In general, the children found the tasks interesting but challenging, particularly the younger NL control children. A number of children were unable to engage with the activities, particularly the PTT. Missing data ranged between 6 and 56% for the NL control group (6% plurals, 9% digit recall 19% CNRep, 56% past tense) and between 0 and 34% for the children with SLI who only had missing data on the PTT. The spread of scores for the children who did complete the PTT ranged from 3 to 77% for the children with SLI and from 3 to 60% for the younger controls. Hence, it appears that the PTT (at least the one used in the present study) may be too difficult for children at this language age.

The means and standard deviations for each group on each of the risk marker tasks are shown in Table 1. Note that the measures used included centile scores for digit recall and CNRep and percent correct for the past tense and plural tasks.

First, a series of ANOVAs was carried out on digit recall, CNRep, past tense and plural tasks comparing children with SLI and younger NL controls. These revealed significant group differences for the two processing-related tasks, i.e. digit recall ($F(1,59) = 34.99$, $p < 0.001$) and CNRep ($F(1,56) = 21.95$, $p < 0.001$). No significant differences between groups were found for the two linguistic tasks, i.e. PTT ($F(1,33) = 0.88$, $p = 0.356$) and plural task ($F(1,60) = 0.04$, $p = 0.839$).

Table 1. Means and standard deviations for the four risk markers for children with SLI and younger NL controls

<table>
<thead>
<tr>
<th>Task</th>
<th>SLI</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit recall (centiles)</td>
<td>mean 28.81</td>
<td>63.10</td>
</tr>
<tr>
<td></td>
<td>SD 21.86</td>
<td>23.42</td>
</tr>
<tr>
<td>CNRep (centiles)</td>
<td>mean 27.59</td>
<td>63.89</td>
</tr>
<tr>
<td></td>
<td>SD 29.70</td>
<td>28.89</td>
</tr>
<tr>
<td>Past tense (% correct)</td>
<td>mean 20.52</td>
<td>14.50</td>
</tr>
<tr>
<td></td>
<td>SD 20.04</td>
<td>16.26</td>
</tr>
<tr>
<td>Plural (% correct)</td>
<td>mean 42.38</td>
<td>43.50</td>
</tr>
<tr>
<td></td>
<td>SD 20.33</td>
<td>23.12</td>
</tr>
</tbody>
</table>
Out of interest, raw score information for the four tasks is presented in appendix A. A series of ANOVAs showed no significant differences between the two groups of children for any of the variables when raw scores were used with the exception of digit recall where children with SLI had a higher mean raw score than the younger NL control children ($F(1,59) = 15.4$, $p < 0.001$). Thus, the data on raw scores revealed no interesting differences between the two groups of children. Given these results, no further analyses using raw scores were carried out (but see the Discussion for further consideration of this issue).

Second, for each risk marker, analyses were completed for sensitivity and specificity. This analysis requires that a threshold score be used as a cut-off for predicting group membership. Cut-off points were selected for each test; children scoring at or below the cut-off were defined as ‘children with impairment’ and those scoring above the cut-off were classified as ‘children without impairment’. Two different thresholds were examined for each risk marker. A ROC curve analysis was carried out first. A ROC curve plots sensitivity against specificity for all possible cut-off points through the whole test score range (Silman 1995). Examination of the ROC curve identified the 25th centile (25% correct in the case of the linguistic tests) as the optimum cut-off point for the four tests (i.e. the test scores giving the highest sensitivity with an acceptably high specificity level). However, the 16th centile (16% correct in the case of the linguistic tests) was also investigated as it is often used for clinical identification of children with SLI.

The sensitivity, specificity and accuracy of each test were calculated using the following formulae:

- **Sensitivity**: number of children with SLI scoring at or below cut-off point/total number of children with SLI (i.e. probability that a child with impairment will be correctly identified by the test = true positive rate).
- **Specificity**: number of NL children scoring above cut-off point/total number of NL children (i.e. probability that a child without impairment will be correctly identified by the test = true negative rate).
- **Accuracy**: number of children with SLI identified correctly (sensitivity) added to the number of NL children identified correctly (specificity)/the total number of children with and without impairment.

Sensitivity, specificity and accuracy calculations are presented in table 2 for each risk marker task.

The two processing tasks, CNRep and digit recall, proved to be the most accurate risk markers overall, particularly at the 25th centile. Nonetheless, the question still arises as to the most suitable measure or combination of measures that can positively identify persistent SLI risk in young learning children with early language delay. To answer this question we carried out a stepwise discriminant analysis of the data with all four risk marker tasks included. In this, group membership (SLI or ‘normal’) was the outcome (dependent) measure, while the data from the two processing and the two linguistic tests were the predictor (independent) measures. The analysis was performed iteratively. At the first step, the predictor measure showing the difference between the groups with the highest level of statistical significance was entered into the discriminant model; at the second step, the predictor showing the highest level of statistical significance (after adjustment for the first measure entered) was added to the model; and so on. The
Table 2. Sensitivity, specificity and accuracy of tasks using thresholds of 25th, 16th centiles/percentages correct for children with SLI and younger NL controls

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25th</td>
<td>16th</td>
<td></td>
</tr>
<tr>
<td>Digit recall</td>
<td>53% (17/32)</td>
<td>90% (26/29)</td>
<td>71% (43/61)</td>
</tr>
<tr>
<td>CNRep</td>
<td>66% (21/32)</td>
<td>85% (22/26)</td>
<td>74% (43/58)</td>
</tr>
<tr>
<td>Past tense</td>
<td>59% (19/32)</td>
<td>89% (23/26)</td>
<td>72% (42/58)</td>
</tr>
<tr>
<td>Plurals</td>
<td>16% (5/32)</td>
<td>77% (23/30)</td>
<td>45% (28/62)</td>
</tr>
<tr>
<td></td>
<td>13% (4/32)</td>
<td>80% (24/30)</td>
<td>45% (28/62)</td>
</tr>
</tbody>
</table>

The only variable that entered the discriminant model was non-word repetition. Neither past tense marking, plural marking nor digit recall entered the model, indicating they were not independent predictors of group membership. Non-word repetition on its own provided an overall predictive ability of 74.1% (sensitivity 72%, specificity 77%).

Discussion

The results suggest that processing tasks involving short-term memory, particularly non-word repetition, have potential as possible risk markers of SLI when assessing young language-learning children. As our interpretation of risk marker does not assume causality or uniqueness, we do not want to emphasize the results with regard to discriminating or categorizing children as SLI or not. Instead, we are suggesting that young children (under 5 years) who already present with a clinical picture of slow language development are potentially at risk for SLI if their performance on non-word repetition falls below the 25th centile, i.e. they fall in the bottom quarter of the normal distribution.

Having said this, there are a number of issues that necessarily indicate that these findings should be regarded as interesting but preliminary and be treated with caution. First, we are not convinced that the linguistic tasks used, particularly the past tense activity, were appropriate enough to engage young NL children. The proportion of missing data for this group (more than half the children) suggests that at least this particular version of the PTT may be too difficult for young language-learning children and therefore we may have not obtained optimum data from them. There is a need to explore the possibility that very young children may not have developed sufficient skills in the area of verb morphology to assess them in this area or to distinguish them from children with SLI consistently. Thus, there
is no doubt that additional research is warranted. In addition, the present task involved the use of line drawings, and future research should consider more appropriate tasks for younger language-learning children perhaps using colour pictures (Rice and Wexler 1996) or real objects (Rice and Wexler 2001).

Second, and very importantly, although we have established that processing tasks have potential as risk markers for SLI, we have done this by using centile information in the processing tasks and percentage correct measures in the linguistic tasks (as normative, centile data were not available for the linguistic tasks). Recall that centiles are anchored to the children’s chronological ages. In other words, we have demonstrated that children with SLI, when compared with their chronological age peers, perform much worse on processing tasks than younger NL-learning children when compared with their age peers. This, it can be argued, is not particularly surprising. We have not demonstrated that the SLI children performed more poorly in processing tasks than the younger NL-learning children in absolute terms. It could be the case that a younger NL child performs worse than most of the children with SLI and yet has a higher centile as the centile is anchored on comparisons with same age peers. Interestingly, when examining the mean raw scores in this study, this was not evident as both groups of children (SLI and NL) had virtually the same mean raw score (see appendix A). It is our opinion that further research is needed in this area to examine in more detail the performance of children with SLI in short-term memory tasks such as non-word repetition. What we have in mind is the type of detailed analysis carried out by Dollaghan and Campbell (1998) with older children which involves detailed transcription of each repeated phoneme in order to ascertain whether children with SLI have qualitatively different productions than NL children when engaged in non-word repetition tasks. At present, the information available is quite crude in that a judgement is made about whether or not the child has repeated a particular item correctly, with little or no information at the phoneme level. We believe that interesting research possibilities lie ahead if much more detailed analyses are carried out in an attempt to discover if there is a certain type of error made by children with SLI in their production of novel words that is not (or very rarely) present in NL development.

Finally, although it needs to be noted that non-word repetition appears to be a powerful risk marker for SLI, it is still unclear conceptually what the impact of a repetition deficit entails. Gathercole and Baddeley (1996) see non-word repetition as closely linked to a range of language skills, including novel word learning, comprehension of spoken language and literacy development. It is not surprising then that phonological processing limitations have been demonstrated in children with a variety of disorders, e.g. those with more global learning disabilities such as Down’s syndrome (Jarrold et al. 2000) and those with severe reading difficulties (Kamhi and Catts 1986), as well children with purer language impairment (Bishop et al. 1996). These findings suggest that difficulties with non-word repetition may be more related to any language impairment and that the specific nature of SLI still remains to be fully understood.

Indeed, much about phonological processing and non-word repetition remains to be clarified. Snowling et al. (2000) have postulated a common deficit of phonological short-term memory underlying language and literacy difficulties but, surprisingly, Kamhi and Catts (1986) found that non-word repetition was not closely linked with other phonological awareness measures (which are themselves closely related to early literacy and which would generally be agreed also to involve considerable
phonological short-term memory processing). Bishop et al. (1996) predicted, but did not find, a link between non-word repetition and vocabulary development; they also reported the persistence of non-word repetition difficulties in children whose language problems have resolved. In contrast, it is interesting to note the impairment of non-word repetition in children with developmental dyslexia but without a history of speech and language delay. Direct comparisons between children with SLI and children with reading difficulties with this issue in mind have not yet been reported and will be an important future research area.

As a weak marker of language delay, a non-word repetition problem may be a necessary but is unlikely to be a sufficient factor in language delay and its impact and manifestation may be influenced by an individual child’s profile of linguistic and cognitive abilities. It may also be particularly important at certain stages of development; in vocabulary learning, at an early stage, when children are heavily reliant on word form and structure and less on semantic links to establish new vocabulary; in reading development, at the alphabetic stage where children are dependent on phoneme-grapheme correspondences.

Given the above considerations, we repeat our mild interpretation of risk marker for SLI, where we suggest that a risk marker is a symptom that in combination with other information, points to increased risk. With this clarification in mind, we believe that the addition of non-word repetition to the assessment of young language-learning children can be fruitful and should be included in clinical practice. In addition, continued research involving non-word repetition and other short-term memory tasks should increasingly provide more information that may refine our understanding of possible risk markers for SLI in young language-learning children.

Acknowledgements
This research was supported by a grant from the Wellcome Trust (Project No. 053456) awarded to G.C.-R. Thanks to Brian Faragher for statistical advice and to Kirsten Windfuhr and Amy Skipp for help with data collection and preliminary analyses. Many thanks also to the schools and families who helped make this research possible.

Appendix A: Means and standard deviations for the four risk markers for children with SLI and younger NL controls (raw scores)

<table>
<thead>
<tr>
<th>Task</th>
<th>SLI</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit recall</td>
<td>mean 10.13</td>
<td>7.03</td>
</tr>
<tr>
<td></td>
<td>SD 2.97</td>
<td>3.19</td>
</tr>
<tr>
<td>CNRep</td>
<td>mean 15.47</td>
<td>15.89</td>
</tr>
<tr>
<td></td>
<td>SD 6.28</td>
<td>8.01</td>
</tr>
<tr>
<td>Past tense</td>
<td>mean 6.19</td>
<td>4.36</td>
</tr>
<tr>
<td></td>
<td>SD 6.01</td>
<td>4.86</td>
</tr>
<tr>
<td>Plural</td>
<td>mean 13.16</td>
<td>13.40</td>
</tr>
<tr>
<td></td>
<td>SD 6.29</td>
<td>7.15</td>
</tr>
</tbody>
</table>
References


