

 Open access • Journal Article • DOI:10.1080/13682820500442644

## Development of the Grammar and Phonology Screening (GAPS) test to assess key markers of specific language and literacy difficulties in young children.

— [Source link](#) 

Hilary Gardner, Karen Froud, Alastair McClelland, Heather K. J. van der Lely

**Institutions:** University of Sheffield, Columbia University, University College London

**Published on:** 10 Sep 2006 - International Journal of Language & Communication Disorders (Taylor & Francis)

**Topics:** Language disorder, Standardized test, Language acquisition, Grammar and Test (assessment)

Related papers:

- [Psycholinguistic markers for specific language impairment \(SLI\).](#)
- [Children with Specific Language Impairment](#)
- [Prevalence of Specific Language Impairment in Kindergarten Children](#)
- [Phonological memory deficits in language disordered children: Is there a causal connection? ☆](#)
- [Domain-specific cognitive systems: insight from Grammatical-SLI.](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/development-of-the-grammar-and-phonology-screening-gaps-test-3ml26g9l1a>

# Development of the Grammar and Phonology Screening (GAPS) test to assess key markers of specific language and literacy difficulties in young children

Hilary Gardner<sup>‡</sup>, Karen Froud<sup>§</sup>, Alastair McClelland<sup>¶</sup> and Heather K. J. van der Lely<sup>†</sup>

<sup>‡</sup>Department of Human Communication Sciences, University of Sheffield, Sheffield, UK

<sup>§</sup>Department of Biobehavioral Science, Teachers College, Columbia University, New York, NY, USA

<sup>¶</sup>Department of Psychology and <sup>†</sup>Centre for Developmental Language Disorders and Cognitive Neuroscience, Department of Human Communication, University College London, London, UK

(Received 5 April 2005; accepted 18 October 2005)

## Abstract

*Background:* Despite a large body of evidence regarding reliable indicators of language deficits in young children, there has not been a standardized, quick screen for language impairment. The Grammar and Phonology Screening (GAPS) test was therefore designed as a short, reliable assessment of young children's language abilities.

*Aims:* GAPS was designed to provide a quick screening test to assess whether pre- and early school entry children have the necessary grammar and pre-reading phonological skills needed for education and social development. This paper reports the theoretical background to the test, the pilot study and reliability, and the standardization.

*Methods:* This 10-min test comprises 11 test sentences and eight test nonsense words for direct imitation and is designed to highlight significant markers of language impairment and reading difficulties. To standardize the GAPS, 668 children aged 3.4–6.6 were tested across the UK, taking into account population distribution and socio-economic status. The test was carried out by a range of health and education professionals as well as by students and carers using only simple, written instructions.

*Results:* GAPS is effective in detecting a range of children in need of further in-depth assessment or monitoring for language difficulties. The results concur

Address correspondence to: Heather K. J. van der Lely, Centre for Developmental Language Disorders and Cognitive Neuroscience, Department of Human Communication Science, University College London, 2, Wakefield Street, London WC1N 1PF, UK; e-mail: h.vanderlely@ucl.ac.uk

with those from much larger epidemiological studies using lengthy testing procedures.

*Conclusions:* The GAPS test (1) provides a successful screening tool; (2) is designed to be administered by professionals and non-professionals alike; and (3) facilitates identification of language impairment or at-risk factors of reading impairment in the early educational years. Thus, the test affords a first step in a process of assessment and targeted intervention to enable children to reach their potential.

*Keywords:* grammar, phonology, language, screening, SLI.

## Introduction

Language and communication are universally recognized as some of the essential building blocks for children's educational, social, and vocational development. Recent government policies in the UK (*The National Literacy Strategy*, Department of Education and Skills 2001; *Speaking and Listening Curriculum*, Quality Curriculum Authority (QCA) 2003) and internationally e.g. in the USA (*Put Reading First*, National Institute for Literacy (NIFL) 2001) have therefore stressed the importance of oral language for children's eventual achievement. Various strategies for the monitoring and enhancement of language skills from an early age have been put in place. Health, education, and social services have recognized that developmental language deficits significantly affect children's potential, but there has not been a simple, effective measure of children's language abilities that could identify at risk children. The simplicity of the Grammar and Phonology Screening (GAPS) test screen means it could be routinely administered by professionals and non-professionals alike, either to large cohorts of children before or at school entry, or to children causing concern in their early school years. Whilst no quick screen could hope to identify all the subtly different forms of specific language impairment (SLI), GAPS goes beyond screening for general levels of language functioning as it is designed to target core grammatical and phonological abilities which are known to be impaired in the majority of children with specific language impairment and/or specific reading difficulties (Gathercole and Adams 1993, van der Lely 1996a, 1997, 1998, Bishop 1997, van der Lely and Stollwerk 1997, Conti-Ramsden 2003, Bishop *et al.* 2004, van der Lely and Battell 2003, Rice 2004). Only a small proportion of children with language impairments in areas of social interaction or forms of lexical-SLI have little or no deficits in syntactic and phonological functioning (Nation *et al.* 2004, van der Lely 2005).

The present paper reports on the development and standardization of the GAPS test, first giving details of the theoretical background to the test design, including the choice of stimuli. Next, the pilot study involving 148 children is described with resultant modifications. The results of its standardization on 668 children across the UK are then given in full followed by a small separate set of results on two groups of children with special needs. One group were children already identified as having SLI whilst the others were referred for a variety of reasons including poor educational progress, behavioural difficulties or other reasons.

### *Background to the test*

Communication through language is important in almost every aspect of modern life. Yet, approximately 7% of children who are otherwise apparently developing

normally have an impairment in language acquisition, termed 'SLI' (Leonard 1998). There is increasing evidence that language development depends on multiple underlying faculties that are genetically distinctly specified (Bishop *et al.* 2005) and the majority of these SLI children have variable deficits in components of grammar (syntax, morphology and phonology) as well as other aspects of language such as vocabulary (van der Lely 2005). In addition, such language deficits can co-occur with low IQ or other cognitive impairments such as Down's syndrome (Norbury *et al.* 2002, Laws and Bishop 2003). The language impairment often persists into adulthood (Leonard 1998) and impacts on various aspects of life, including social and emotional well-being.

Many children with SLI also have specific reading difficulties, sometimes in the form of very slow literacy development where syntactic and morphological problems are associated. However many children have specific reading difficulties (dyslexia) where phonological impairment is considered to be core to their deficit (Snowling 2000, Ramus 2003) and these children may or may not have associated language impairment *per se*.

Based on the numbers of children in the UK between 3 and 6 years (3.015 million, in 2000) the 7% prevalence of SLI represents over 211 000 children (Office of National Statistics 2005). Snowling (2000) states that 4% of all school-age children suffer from severe dyslexia and a further 6% from mild-to-moderate dyslexia. Thus, these two disorders make up, by far, the most prevalent disorders in childhood. A screening test identifying such prevalent language difficulties or the antecedents to literacy difficulties, before the child's entrance into formal education would therefore provide a valuable tool. The GAPS test is designed to provide just such a tool to identify children with or at risk for developmental language and/or literacy disorders early in their educational career. It provides a check that children are 'school ready' in key language areas, namely grammar (morpho-syntax) and phonology, specific aspects of language often left untapped by standardized tests for young children. Identifying children who might have undetected grammatical or phonological deficits is clearly advantageous, as it will facilitate appropriate referral for detailed assessment, early diagnosis and intervention, for instance by a speech and language therapist, thereby helping to limit the impact of associated literacy and learning problems. As effective intervention relies on etiological insight it is also important that possible indicators of reading difficulties are identified and distinguished, both those that could arise as part of SLI and those that indicate a predisposition to phonological developmental dyslexia. It is only with insight into initial language function that progress made by typically developing (TD) children in language and literacy skills can be effectively screened, allowing realistic evaluation of the efficacy of relevant educational and other intervention methods.

Grammatical deficits in children have been shown to be resistant to treatment, either through an enriched environment (Newman 1996) or more specific therapy (Nelson *et al.* 1996, Fey *et al.* 1997, Law *et al.* 2003). Other areas of language (social interaction, vocabulary development) typically show greater positive effects (Haynes and Naidoo 1991, Law *et al.* 2003). However there is growing evidence that specialized intervention programmes, targeted with knowledge of children's underlying deficits, such as that reported by Ebbels and van der Lely (2001) and Ebbels (2005), can be effective. The evaluation of the type of intervention required is directly relevant to the aspect of language being measured and the nature of the deficit. A screening test that targets specific, structural aspects of language

functioning is vital as part of the process of identifying subgroups of children who may require such targeted language interventions, as well as those established reading interventions targeting phonological knowledge.

### *Development of the GAPS test*

#### *Critical markers of specific language impairment and specific reading difficulties*

SLI is a heterogeneous disorder that variably affects the different components of language. There is a large body of research showing that children with SLI inconsistently manipulate core aspects of syntax, including tense-marking, assigning thematic roles in passive sentences and embedded phrases and clauses, assigning reference with reflexives and forming questions (for a review, see van der Lely 2004, 2005). The problems with syntactic and morphological aspects of tense-marking and other inflectional forms are a well-established phenotypic characteristic of SLI (Rice and Wexler 1996, Clahsen *et al.* 1997, Leonard 1998, Norbury *et al.* 2001, Conti-Ramsden and Hesketh 2003, van der Lely and Battell 2003, van der Lely 2005). Further, Rice and Wexler (1996) argue that tense-marking provides a good clinical marker of SLI. However, recent research highlights the breadth of the impairment within the grammatical system in not only the Grammatical (G-SLI) subgroup (van der Lely 2005) but in many forms of SLI in children as evinced by replications of the findings summarized above, in different research groups, in both English-speaking children with SLI (Precious and Conti-Ramsden 1988, O'Hara and Johnston 1997, Bishop *et al.* 2000) and in other languages such as German, Greek and Hebrew (Hamann *et al.* 1998, Stavrakaki 2001, 2002, Friedmann and Novogrodsky 2004).

According to the Computational Grammatical Complexity (CGC) hypothesis, children with G-SLI are impaired when performing the mental computations which underlie hierarchical, structurally complex syntactic, morphological or phonological forms (van der Lely 2005). The CGC hypothesis emphasizes the distinction between syntax, morphology and phonology, and their independent and differential effects on sentence processing and production. These three components of grammar show types of computational complexity not found in the processing needed in other cognitive domains (Chomsky 1986, Hauser *et al.* 2002). Performance in syntax, morphology and phonology have also been found to dissociate from other cognitive functions in developmental disorders (Pinker 1994, Clahsen and Almazan 1998, van der Lely *et al.* 1998, Ramus 2003, van der Lely 2005).

The syntactic deficits of children with SLI appear to be restricted to syntactic dependencies that in Chomsky's terms involve 'movement' (Chomsky 1995). Thus, not only tense-marking, but also many other structures involving syntactic dependencies are predicted to be problematic for children with language impairments. Such structures are reflected in the choice of test items given below. They include passive sentences (*The dog is licked by the cat*), and pronominal reference (*The cat is washing herself*) (for a full explanation of these 'syntactic dependencies' and the CGC, a development of Representational Deficit for Dependent Relations (RDDR) hypothesis, see van der Lely 1998, 2005, and van der Lely and Battell 2003). Morphological complexity can be understood with respect to Pinker's Words and Rules model (Pinker 1999) whereby normal developers store irregular forms whole but compute morphologically regular forms using a symbolic rule (*roll + ed*). The CGC and other theories predict that for children with a morphological (rule) deficit,

regularly inflected verbs might be particularly problematic as they are stored in the same way as irregulars and are therefore subject to the same word effects (Marshall and van der Lely 2005). Thus, past tense regular forms are included in the test items.

However, tests of phonological abilities are crucial in differentiating children with language impairment with and without phonological deficit and from those children with only phonological deficit who are at risk for more specific reading difficulties: Whereas many, but not all, children with SLI fail on both tests of morpho-syntax and phonology (Conti-Ramsden 2003, Marshall *et al.* 2003, Gallon *et al.* 2005), children and adults with a phonological form of dyslexia arguably only fail on tests of phonology (Snowling 2000, Ramus 2003). Thus, tests of both morpho-syntax and phonology are needed to provide preliminary indicators of language component impairment and to lay the foundations for appropriate full assessment and remediation.

Non-word repetition forms the second subtest of the GAPS. Here, the test items were designed to reflect increasing phonological complexity that can be found in English words. SLI children's problems with repetition of non-words are well known. One explanation for this problem is that it reflects a deficit in phonological short term memory (P-STM) (Gathercole and Baddeley 1990, Conti-Ramsden 2003). However, both processing and memory are significantly affected by underlying linguistic (phonological) representations (Morgan and Demuth 1996, Nääätänen *et al.* 1997). Thus, more recently, it has been recognized that non-word repetition also reflects phonological abilities, and, moreover, is a good marker of SLI and dyslexia (Bishop *et al.* 1996, Conti-Ramsden and Hesketh 2003, Marshall *et al.* 2003, Gallon *et al.* 2005). Furthermore, Roy and Chiat (2004) showed that non-word repetition tasks can be effectively used with children as young as 2 years of age to discriminate patterns of performance and map change over time.

The hierarchical structural organization of sounds into words is common to all languages (Harris 1994). So-called 'parameters' regulate segmental syllable structure (Onset, Rhyme, Nucleus) and the prosodic-'metrical' word/foot structure, which affects the stress pattern of syllables in words (Marshall *et al.* 2003, van der Lely 2005). These syllable and metrical structures can be either 'marked' (atypical forms) or 'unmarked' (the basic, common language forms) (Chomsky and Halle 1968, Gallon *et al.* 2005, van der Lely 2005). Marked structures are not attested in all languages, and occur later in language acquisition and are therefore deemed to be more complex. Furthermore, marked structures can occur in combination; with the greater the number of marked structures the more complex the word or non-word. Many SLI children show increasing errors in non-word word repetition as the number of marked parameters increases (van der Lely 2004, van der Lely *et al.* 2004). Previous research shows that even short but phonologically complex non-words can cause difficulties (Marshall *et al.* 2002, Roy and Chiat 2004, van der Lely *et al.* 2004). Thus, non-words such as *badrep* are predicted to be harder than *drepa*, as phonologically in *badrep* the initial weak syllable makes it structurally more complex. We exploit this finding in selecting the test items for the GAPS phonology subtest.

Over the past 10 years or more van der Lely and colleagues have developed a number of tests and experimental procedures tapping the three core components of grammar that appear to be core deficits in SLI children (van der Lely 2005). These tests, such as the Verb Agreement and Tense Test (VATT) (van der Lely 1999) the Test of Active and Passive Sentences (TAPS) (van der Lely 1996b), and the

Advanced Syntactic Test of Pronominal Reference (A-STOP) (van der Lely 1997), which assesses pronominal and anaphoric reference in sentences, and the Test of Phonological Structure (TOPhS) (van der Lely and Harris 1999) provide a basis for the choice of sentence and non-word stimuli selected for this screening test — the GAPS test. The original, lengthy tests and procedures have been shown to target exactly those aspects of language that children with language impairment or children at risk for phonological developmental dyslexia fail (van der Lely *et al.* 2004, Ebbels 2005, van der Lely 2005). They have proved to be highly sensitive for detecting language impairment in children when it occurs in isolation (van der Lely 1996a, van der Lely and Stollwerk 1997, van der Lely and Battell 2003), when it co-occurs with low non-verbal IQ (Bishop *et al.* 2000, Norbury *et al.* 2001, 2002), and in distinguishing impaired versus normal grammatical abilities in children with other developmental disorders such as Down's syndrome or William's syndrome, respectively (Clahsen and Almazan 1998, Ring and Clahsen 2003).

In summary, our theoretical framework, alongside previous research findings, provide the background to us targeting pertinent linguistic aspects to be used in the screening test. It should be noted that many available standardized tests do not aim to assess specific aspects of grammar or phonological development. For example, in the Reynell Developmental Language Scales (Edwards *et al.* 1997), a thorough test of general language function, it is possible with single word vocabulary, auditory memory and general world knowledge to achieve correct scores without using grammatical knowledge. In contrast, in the GAPS test, the focus is on arguably the 'core' deficits found in key morpho-syntactic and phonological structures which normally developing children master with no difficulty by around 3–4 years of age but which cause particular difficulties for children with language impairment and those at risk for specific reading difficulties.

### **Aims**

Detailed grammatical assessments of children traditionally require professionally trained people to administer them and interpret the results. Widely available standardized language tests are lengthy procedures, when most clinicians, psychologists, and educationalists are under severe time constraints. This results in only those children who are already exhibiting problems (often severe) being assessed. Alternatively, 'parental concern' has been used as a criterion for referral for further assessment and has been shown to be a more reliable predictor for later language impairment than some standardized tests (Dale *et al.* 2003). However, evaluation of this method reveals over-referral of children without language problems and under-referral of those with deficits (Laing *et al.* 2002). Heath and Hogben (2004) advocate cost effective screening for early literacy difficulties including oral language and phonological skills, but as yet there has not been an available screening test, standardized in the UK, that is easy to administer, therefore unlikely to require significant professional training and that, crucially, is sensitive to identifying language disorders and children at risk for phonological developmental dyslexia. The 'GAPS' test aims to fill this health and educational need. Other research, such as that by Pickstone *et al.* (2002), clearly suggests a role for paraprofessionals in the screening of language development in the preschool population, especially as they could achieve high levels of coverage in a given

community. In addition, school teaching assistants are being encouraged to assume more responsibility for collaborative work with speech and language therapists and other outside agencies (Law *et al.* 2003). Thus, the GAPS test aims to provide a useful tool for such paraprofessionals and professionals alike and also takes a step towards including non-professionals in this process.

## **Methods and procedures**

### *Testers*

A high proportion of testing was carried out by professional and non-professional volunteers in addition to that done by the research team (who mainly carried out testing at the pilot stage). For the standardization, a call for volunteers was made via a professional website covering those working with children in education and health, as well as contact being made through heads of service and professional contacts. The response was overwhelmingly positive. In line with the aims of this screening test, at the standardization stage a high proportion of the testing was carried out by staff and carers not traditionally trained in language assessment. Of the 60 testers (table 4), 18 were teaching assistants, nursery nurses, play-group leaders and mothers who tested 220/720 children. Most were recruited through a lead clinician or educationalist with whom the research team first made contact. 21 testers were speech and language therapists who tested 154 subjects, 14 teachers tested 223 children and finally, seven volunteer undergraduate and postgraduate students of speech/language therapy or psychology tested 123 subjects. Only a small group of 11 children were tested by a member of the research team. Testing took place predominantly in schools, nurseries or the children's homes. The testers were given a manual that contained clear instructions and examples of child responses. Some support was offered through phone and e-mail when required at the beginning of testing but in no instance was further help requested. No interrater reliability between professional and other testers took place but a feasibility assessment of a small sample of children were marked independently on the tick/cross system by a paraprofessional tester and speech and language therapist researcher, with 95–100% agreement, at an early stage of the standardization.

### *Subjects*

The pilot subjects were gathered in three stages as development of the test progressed. Initially 54 TD children were recruited between the ages of 4.10 and 8.11 years. These children all came from Central London. The second stage involved more TD subjects and 17 language impaired children (analysed separately) from speech and language therapists in London and Leeds. The age range here was 3.01–8.01 years but this age range was reduced once the pattern of response had been analysed and a ceiling effect identified in the TD children. The third stage involved additional TD children from Greater London, Leeds and Arbroath, Scotland, where the main testers were based. The final usable cohort totalled 148. The schools who volunteered subjects received a small sum towards their school fund or a charity of their choice. Although there was no intention to collect a balanced sample at this stage socio-economic status information by occupation was collected on 68 of the



subjects to ensure all levels were at least represented. (The data was not requested on the subjects at the first stage.) An abbreviated version of the Standard Occupational Classification (Office of National Statistics 2000a, b) in which the nine major classifications were reduced to four categories was used. There was a greater representation of children from parents with professional, managerial and technical occupations than unskilled, the numbers being as follows. Of 68 children, 22 fell in the managerial and professional classes, 17 were in higher technical, 16 were in skilled manual and non-manual occupations, seven in elementary occupations and 'other'.

Subjects for the standardization were gathered from across the UK. Full details of the cohort are given in the Results. Every effort was made to gather a cohort balanced for socio-economic status and rural/urban populations as far as this was possible, given the restrictions of access to volunteer testers and parental consent for testing. Consent was obtained in several areas from whole classes or preschool groups.

Two small groups of children either causing concern at school or already diagnosed with SLI are reported separately. The former (32 children) were those who were being seen or were about to be assessed by specialist external school support teaching teams, having been referred by their schools because of a variety of concerns. Further details were not always stated explicitly but they included poor behaviour, specific or general poor attainment in school and so forth. There was no immediate implication that these children had either SLI or reading difficulties but the school support team were keen to try the test out on a population already showing reason for concern. The other small group of 17 children diagnosed as SLI were all from language resourced based schools or units where teachers or SLTs had volunteered to assess them on GAPS.

#### *Test procedures: elicited imitation*

An elicited imitation procedure is simple to administer and yet allows more control over administration and analysis than other procedures (McDaniel *et al.* 1996, Crain and Thornton 1998, Lust *et al.* 1996). It was therefore chosen as the basis for the GAPS test, especially as some paraprofessionals and non-professionals, unused to formal test procedures were to take part in the administration. The test comprises two elicited imitation tasks: one to test morpho-syntax, involving repetition of sentences, and one to test phonology involving repetition of non-words. Children are shown a picture book and asked to repeat a story to an alien figure called 'Bik', who only understands when children speak to him. The story is told to them by the tester, one sentence at a time for the children to repeat to Bik. The child is then asked to repeat the small set of non-words which are deemed to be in the alien's language.

This methodology allows a focus on specific aspects of grammar and phonology, which can be precisely manipulated. Since the 1950s (Brown 1957), elicited imitation has proved to be a highly reliable, powerful and valid method for assessing core grammatical knowledge (syntax, morphology, phonology) (Thornton 1995, Crain and Thornton 1998, Lust *et al.* 1996). Further, it reveals the child's own grammatical ability rather than assuming that of the adult and minimizes confounds with other non-linguistic cognitive factors (Lust *et al.* 1996). Conti-Ramsden and colleagues

## *GAPS test to assess key speech markers in young children*

found elicited sentences to be the most reliable psycholinguistic marker of language impairment, with high levels of sensitivity and specificity. Furthermore, this procedure correctly identified the majority of children whose current language status fell into the normal range, despite a history of SLI (Conti-Ramsden *et al.* 2001).

### *Test stimuli*

#### *Section 1: Test of grammar (morpho-syntax)*

Sixteen sentences were presented: two practice items, 11 test sentences and three simple declarative filler sentences. These assessed subject-verb agreement, tense-marking (past, future) e.g. ‘The cat wanted some milk’, phrasal embedding, e.g. ‘the cat with the bell is happy’, dative construction, e.g. ‘The dog gives the cat the milk’, object question formation, reversible passive construction and anaphoric and pronominal reference. The choice of the sentences was based on our previously developed tests and procedures as reported above.

Finally, careful control of all vocabulary items was employed: all words have an early age of acquisition (e.g. *cat*, *dog*) and are familiar to children regardless of socio-economic or cultural variation. In addition, only words with a simple phonological structure are included. Thus, the test minimizes the likelihood of failure due to the subjects not knowing the words or being able to pronounce them.

#### *Section 2: Test of phonology*

Ten non-words (two practice non-words and eight test non-words) were a selected subset of items from the non-word repetition test of phonological structure (TOPhS) (van der Lely and Harris 1999, Gallon *et al.* 2005), where prosodic structure is carefully manipulated. Specifically, the selected non-words vary in complexity on the following parameters: (1) marked Onset, Rhyme and final nucleus (e.g. dremp), (2) marked Rhyme, with an initial unstressed syllable (Iambic stress pattern, rather than a strong–weak trochaic stress pattern (e.g. bademper), (3) marked Rhyme and final Nucleus with Iambic structure (e.g. difimp), and (4) marked Onset and Rhyme with Iambic structure (weak–strong–weak stress pattern) (e.g. padrepper). Previous research revealed that a sample of 40 normally developing children between 4;6 and 6 years produced non-words with these structures correctly between 85 and 100% of the time (Gallon *et al.* 2005). All items were limited to three syllables or less in order to minimize the effect of phonological short term memory which is critical at four syllables or more (Gathercole and Baddeley 1990).

### *Modifications to the test following the pilot study*

At the pilot stage, the GAPS test was administered to 148 children, as described above. Minor modifications to the test procedure were consequently made after the third stage of the pilot was complete. First, practice items were added with an additional ‘filler’ sentence at the outset. Second, a lead-in sentence was added to the items where children were required to repeat a question, in order to avoid very young children simply answering the question. Third, modifications were made to

the test administration manual and examples of potentially difficult situations that could occur were added. No actual test items were changed between the pilot and final versions of the test. However, because of these slight modifications, the results from the pilot study were not included in the larger standardization. As mentioned before, the lower age limit of the test was raised from 3;0 to 3;4 years as many younger children were unable to proceed, and second, the upper limit was reduced from 8 to 6;5 years due to a clear ceiling effect.

### *Administration*

The test comprises a short picture storybook about a cat and a dog, which serves as prompts for the child to repeat the short sentences (section 1) and made-up words (section 2). The test takes on average 7 min (range 5–10 min) to administer. Little or no training was offered to testers, but a simple manual with clear instructions and examples of scoring was provided. The testers were asked for comments about the test and its ease of administration at each stage of development. Following the pilot study, examples of most types of error or eventuality were included in the manual for reference, for example dealing with issues of accent to ensure children are not penalized and ensuring only key structures are marked as correct or incorrect.

### *Presentation*

#### *Section 1*

The administrator shows the child the picture booklet and tells the child that he is going to tell a little story. Then the administrator opens the booklet to the first page and introduces a stand-up alien ‘Bik’:

Administrator: This is Bik. He only understands when children speak to him, and he would like to hear the story too. So, when I say something, you say it to Bik. Listen really carefully and make sure you say everything to Bik just the same as I say to you. Okay? Let’s Practice. Hello Bik. Here is the story.

The administrator waits for the child to repeat the practice sentence. If the child fails to repeat the sentence, the administrator says: ‘Go on. You say the same thing to Bik’. If the child repeats the two practice sentences satisfactorily the administrator starts to present the test sentences. Where a child failed to repeat the practice items then they were not included in the standardization cohort. If they ceased attempting to repeat the test items further into the test then their score was counted, even if the result was zero.

#### *Section 2*

For the non-word repetition task, the administrator explains that it is the child’s turn to repeat some of Bik’s words. Two simple practice items are spoken by the administrator before the eight test items. Note: regional vowel variation is not targeted, so does not comprise a problem. (This is mentioned in the manual.) If the child attempted the practice items then their score was recorded from then on, as above. One repetition of a non-word was allowed if the child failed to respond.

### *Scoring*

Details were kept as simple as possible. The test form uses a simple tick/cross response. Within the test sentence, the crucial words or parts of words to be scored as correct/incorrect are highlighted in bold on the administration-response form. In this way, exactly those pertinent aspects of morpho-syntactic abilities are assessed, relatively independently of memory failure, vocabulary knowledge etc. For section 2, the test of phonology, accurate repetition of the non-word is marked as correct and any mistakes are marked as incorrect. No transcription was required. The procedure for non-responses is as outlined above with any initial attempt at an actual test item scored even if the child subsequently failed to complete the test. Any child who did not get past the practice items was removed from the cohort.

A table (table 6) resulting from the standardization of the test provides conversion of raw scores into percentile scores relative to the child's age. Children falling below the fifth, tenth or 15th percentile are identified. We recommend that any child falling below the tenth percentile is referred for further assessment, and those falling between the tenth and 15th percentiles are retested after 6 months. However, before reporting the results of the full standardization of this test we briefly report the pilot study, which assessed the reliability and validity of the GAPS.

## **Results and outcomes**

### *Pilot study*

#### *Validity*

To assess the validity of the GAPS in correctly identifying those children with language or pre-reading phonological deficits, three standardized tests of language functioning were administered by researchers and SLTs to the same 148 children who were assessed during the pilot study. These tests were the British Picture Vocabulary Scale — Revised (BPVS) (Dunn *et al.* 1997) which assesses comprehension of single word vocabulary; two subtests from the Clinical Evaluation of Language Fundamentals — Pre-School (CELF) sentence structure and word structure sub-tests, which assess sentence understanding and expressive morpho-syntactic abilities respectively (Wiig *et al.* 2000); and The Children's Test of Nonword Repetition (CN-Rep) (Gathercole and Baddeley 1996). The CN-Rep was administered to children of 4 years upwards due to its standardization range. It should be noted, however, that whereas subtests on GAPS are specifically designed to pick up children with grammatical or phonological deficits, these tests tap other components of language (e.g. vocabulary) or assess a wide range of grammatical structures, some of which are different to those described in the research literature as significant clinical markers of SLI. The exception is the CN-Rep, which although designed primarily to tap phonological-short-term memory, also taps phonological knowledge. Therefore, we predicted that generally there should be a moderate rather than high positive correlation between the GAPS and these previously standardized tests.

Partial correlations, controlling for age were carried out to assess the relations between the two GAPS sub-tests and the above standardized language measures. The GAPS grammar test (sentence repetition) section showed highly significant correlations at the  $p < 0.001$  level with all four tests (BPVS  $r(142) = 0.440$ ; CELF

Sentence Structure  $r(142)=0.524$ ; CELF word structure  $r(142)=0.427$ ; CN-Rep  $r(112)=0.579$ ). Thus, these correlations showed that between 18 and 34% of the variance was being accounted for. The correlations were lower, but nonetheless significant between the GAPS phonology test (non-word repetition) and the BPVS  $r(140)=0.257$ ,  $p=0.002$ ; CELF sentence structure  $r(140)=0.334$ ,  $p=0.001$ ; and the CELF word structure  $r(140)=0.221$ ,  $p=0.008$ , and accounted for between 5 and 11% of the variance. This is clearly expected due to the very different language components that are being tapped by the GAPS phonological subtest and the BPVS and CELF tests. In contrast, a highly and very strong correlation was found between the GAPS phonology (non-word repetition) test and the CN-Rep ( $r(107)=0.671$ ,  $p<0.001$ ). Thus, the GAPS test shows the strongest correlations with those standardized language tests that are more likely to be tapping similar language components with respect to grammar or phonology and generally shows significant correlations with these longer tests of language, even when age is partialled out.

### *Reliability*

The internal consistency of the pilot test was measured by computing Cronbach's alpha for each component of the test and for the test overall. For the sentence repetition component  $\alpha=0.858$ , and for the non-word repetition component,  $\alpha=0.729$ . These values indicate that the test has good/very good internal consistency. Further, for both components all items were positively correlated with a scale composed of the remaining items and the removal of a particular item led to a reduction in the value of Cronbach's alpha in 36 of 38 cases (the alpha was unchanged in the remaining two cases).

### *Standardization*

For the standardization sample, the test was administered to a minimum of 75 and a maximum of 150 children in each of six age bands from 3;4 to 6;8 years. The final usable sample had a total of 668 children. The sample was carefully targeted to control, as far as possible, for regional location (both rural and urban areas were included) and socio-economic status. The latter was determined using the full classification of parental occupation according to the Office of National Statistics (2000a, b) coding index for the UK. A proportion of children from ethnic minorities was included but only those for whom English was the first language.

### *Characteristics of the standardization sample*

The age and gender distributions of the standardization sample are presented in table 1. Gender information was available for 94% ( $n=628$ ) of the total sample ( $n=668$ ). Ethnicity information was available for 96% ( $n=643$ ) of the sample. Using a broad classification criterion, 94.3% of the sample was White, 3.1% Asian and 2.6% Black/other. The geographical distribution of the standardization sample is presented in table 2. Data for the population were obtained from the Office of National Statistics (2005).

Anonymous marking of score sheets by some testers resulted in missing information on gender.

**Table 1. Distribution of the standardization sample broken down by age and gender**

Age range	Female	Male	Missing	Total
3;4–3;11	31	32	23	86
4;0–4;5	40	61	7	108
4;6–4;11	68	82	0	150
5;0–5;5	65	81	0	146
5;6–5;11	53	49	0	102
6;0–6;8	29	37	10	76
Total	286	342	40	668

There was a statistically significant difference between the expected and observed frequencies ( $\chi^2(10)=299.3, p<0.001$ ). Table 2 reveals that certain areas of the country were over-represented (e.g. Yorkshire and Humberside, the East and Scotland) and other areas underrepresented (e.g. the North, the Midlands, the South West and Northern Ireland). There was a small but significant effect of geographical location on test performance (after controlling for age): Sentence repetition,  $F(10, 656)=5.05, p<0.001, \eta_p^2=0.071$ ; non-word repetition,  $F(10, 656)=2.17, p<0.05, \eta_p^2=0.032$ . Thus, 7.1% of the variability in sentence repetition scores was attributable to geographical location, and 3.2% of the variability in non-word repetition scores. Thus, geographical location accounts for only a small amount of the variability in our sample. Information regarding parental occupation was available for 73.1% of the standardization sample ( $n=448$ ). For 120 of the subjects, the parental occupation was unavailable due to personal information being withheld or omitted. Occupational category was determined using the Standard Occupational Classification (Office of National Statistics 2000a, b) which has nine major classifications (table 3).

Again, there was a statistically significant difference between the expected and observed frequencies ( $\chi^2(8)=231.4, p<0.001$ ). In particular, administrative and secretarial, and sales and customer services were underrepresented, and most notably elementary occupations were over-represented. However, unlike geographical location, test performance was unrelated to the occupational background of

**Table 2. Geographical distribution of the standardization sample**

Location	Percentage in the population	Percentage in the sample	Expected number in the sample	Observed number in the sample
North	15.6	10.2	104	68
Yorkshire and Humberside	8.5	21.6	57	144
East Midlands	7.0	2.7	47	18
West Midlands	9.1	3.1	61	21
East London	9.2	15.4	61	103
London	12.9	11.1	86	74
South East	13.6	13.3	91	89
South West	7.8	2.4	52	16
Scotland	8.1	14.7	54	98
Wales	4.9	5.4	33	36
Northern Ireland	3.3	0.1	22	1
	100.0	100.0	668	668

**Table 3. Distribution of occupational groups for the standardization sample**

Occupational category	Percentage in the population	Percentage in the sample	Expected number in the sample	Observed number in the sample
Managers and senior officials	14.9	12.5	67	56
Professional occupations	11.2	12.7	50	57
Associate professional and technical	13.9	13.8	62	62
Administrative and secretarial	13.4	4.0	60	18
Skilled trades	11.2	12.7	50	57
Personal service	7.0	4.9	31	22
Sales and customer service	7.7	3.6	34	16
Process, plant and machine operatives	8.7	9.8	39	44
Elementary occupations	11.9	34.8	55	156
	100.0	100.0	448	448

the parents (with age of child controlled): Sentence repetition,  $F < 1$ ,  $\eta_p^2 = 0.008$ ; non-word repetition,  $F(9, 477) = 1.22$ ,  $p = 0.282$ ,  $\eta_p^2 = 0.022$ .

*Possible effect of the type of tester on scores*

Table 4 provides details of the mean scores, adjusted for the age of the child, for each tester type.

There was a small but significant effect of tester occupation on test performance (after controlling for age): Sentence repetition,  $F(3, 663) = 7.19$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.032$ ; non-word repetition,  $F(3, 663) = 8.15$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.036$ , overall,  $F(3, 633) = 8.75$ ,  $\eta_p^2 = 0.038$ . Thus, 3.2% of the variability in sentence repetition scores, 3.6% of the variability in non-word repetition scores, and 3.8% of the variability in overall performance was attributable to the tester's occupation. Thus, tester occupation accounts for only a small amount of the variability in the data. It is noteworthy that the nursery nurses/teaching assistants and other testers with little or no experience in language assessments gave the highest scores on both the sentence repetition and non-word repetition tasks but it is also interesting to note that the least variability is found in this group. The greatest variability in the scores is found in the group of speech/language therapists.

*Derivation of the norms*

Norms were derived separately for the two sections of the test (sentence repetition and non-word repetition). Age bands were constructed to achieve a reasonably linear increase in performance on both components with increasing age. The mean number of correct repetitions for the sentence task and the non-word task are presented in figures 1 and 2, respectively.

**Table 4. Adjusted means (SD) on both subtests as a function of the type of tester**

Tester occupation	Sentences	Non-words
Nursery nurse ( $n = 175$ )	9.33 (2.43)	6.37 (1.71)
Researcher/student ( $n = 134$ )	9.17 (2.97)	5.40 (1.93)
Speech therapist ( $n = 180$ )	8.29 (2.83)	5.62 (2.44)
Teacher ( $n = 179$ )	8.52 (2.67)	5.61 (2.00)

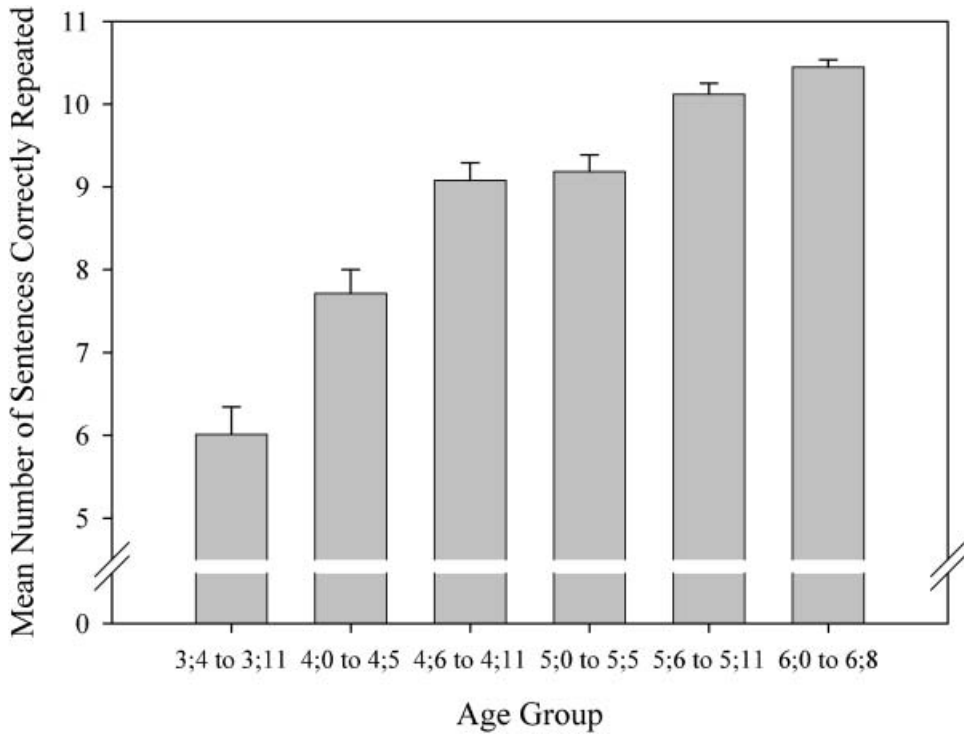


Figure 1. Mean number of sentences correct as a function of age group.

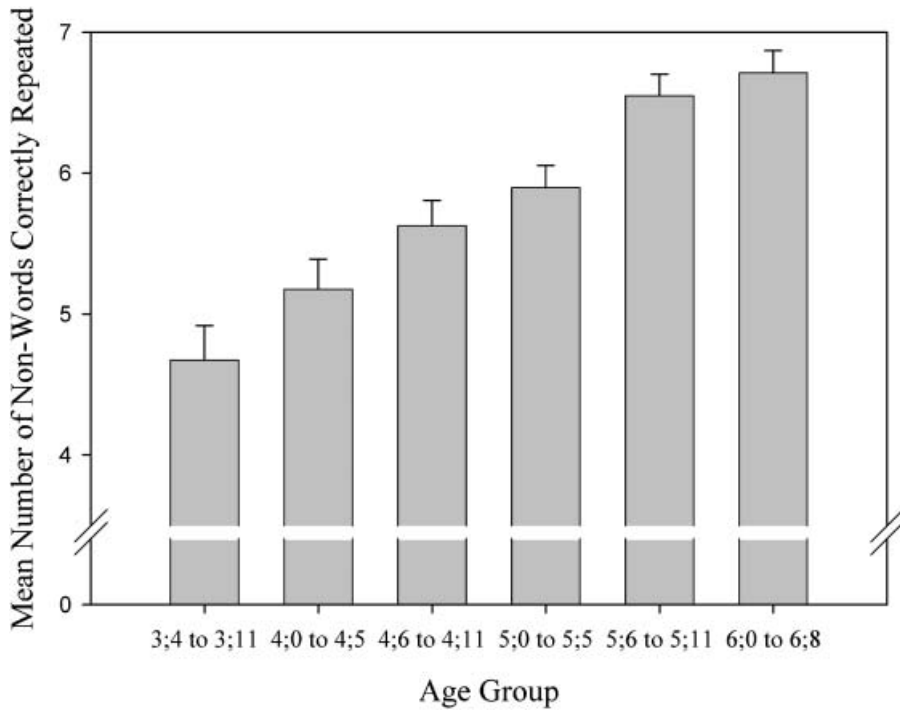


Figure 2. Mean number of non-words correct as a function of age group.



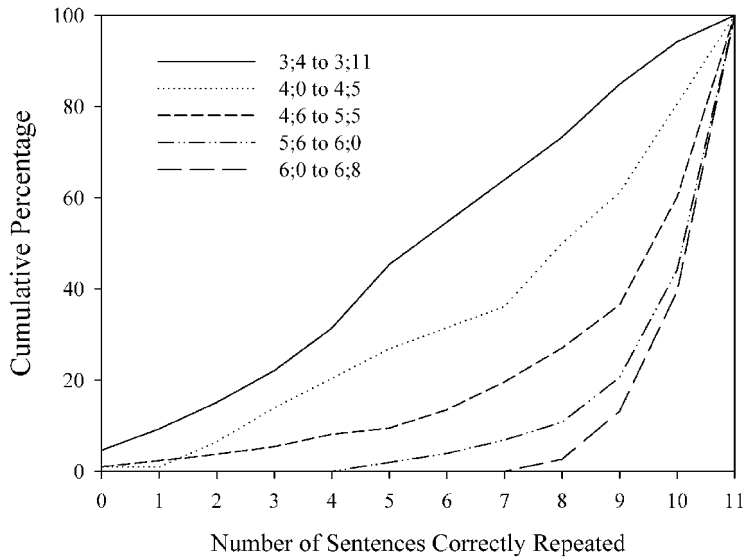


Figure 3. Cumulative percentage of children correctly repeating a given number of sentences within each age band (new groupings).

It is evident from inspection of the bar-charts presented in figures 1 and 2 and cumulative percentages derived from these figures that the age groups 4;6–4;11 and 5;0–5;5 are not well differentiated on either of the test components (particularly sentence repetition). Statistical analysis confirmed this observation. There was no significant difference in mean performance between these two groups on any of the measures (sentences;  $t(294)=0.36$ ,  $p=0.721$ ; non-words;  $t(294)=1.13$ ,  $p=0.260$ ; overall;  $t(294)=0.79$ ,  $p=0.431$ ). As a consequence of these findings it was decided to combine these two age groups to form a single group (4;6–5;5).

Figure 3 presents the cumulative percentage of children who correctly repeated a given number of sentences with the middle two age groups combined.

Figure 4 presents the cumulative percentage of children who correctly repeated a given number non-words with the middle two age groups combined.

Once the new groupings had been established mean performance on both subtests were calculated as a function of age.

There was a significant effect of age group for both tasks (sentences:  $F(4, 618)=47.53$ ,  $p<0.001$ ; non-words:  $F(4, 618)=15.56$ ,  $p<0.001$ ) but no significant effect of gender (sentences:  $F(1, 618)=3.02$ ,  $p=0.08$ ; non-words:  $F<1$ ). There was no interaction between age group and gender (sentences:  $F<1$ ; non-words:  $F(1, 618)=1.43$ ,  $p=0.224$ ). Planned comparisons between the five age groups revealed that all differences were significant with the exception of 5;6 to 5;11 versus 6;0 to 6;8 (sentences: mean=10.12 versus 10.47,  $t(618)=0.93$ ,  $p=0.18$ ; and non-words: mean=6.55 versus 6.75,  $t(618)=0.62$ ,  $p=0.27$ ). Although the test means do not distinguish clearly between children in the two upper age groups, the cumulative scores differentiate the groups. Thus, we considered that this warranted maintaining these two separate age groups, between 5;6 and 6;8 years.

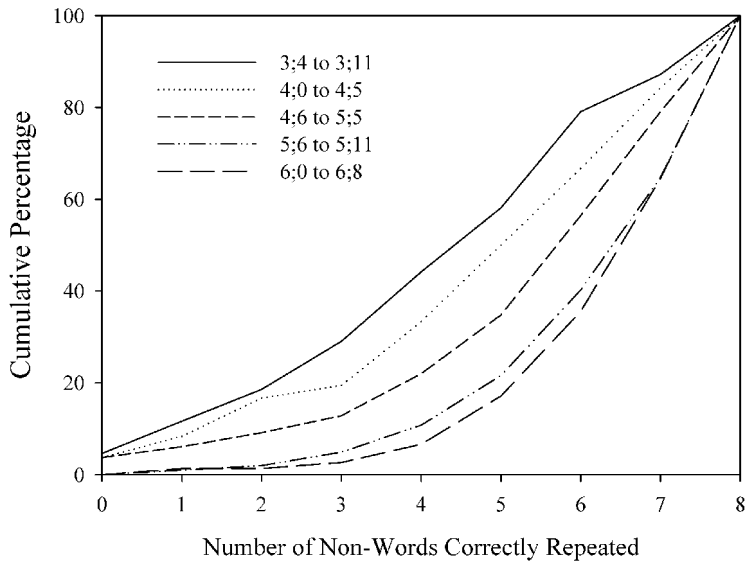


Figure 4. Cumulative percentage of children correctly repeating a given number of non-words within each age band (new groupings).

*Correlation between performance on two subtests*

The correlations between performance on the sentence repetition and non-word repetition components of the test as a function of age group was significant at  $p < 0.01$ , with the correlations varying between 0.68 and 0.41 for the majority of age groups, with the lower correlations at the older ages. The correlation was lower ( $r = 0.28$ ,  $p < 0.01$ ) for the 5;6–5;11 age group. One possible reason for this could be the impact of learning to read at this age and starting formal education. In addition, at the upper age range, scores were approaching ceiling, so the relations between sentence repetition and non-word repetition reduces in the older age groups. Note, however, the raw scores continue to differentiate the groups and provide the means of screening children’s performance.

Based on the revised cumulative distributions, tables 6 and 7 were constructed. Each table gives the percentage of children within each age group who obtained a particular score or lower on each component of the test. An individual child’s score can be located in the left hand column, and by entering the appropriate age group,

**Table 5. Mean (SD) performance on the sentence repetition task and the non-word repetition task as a function of age (new groupings)**

Age group	Repetition task	
	Sentences	Non-words
3;4–3;11	5.7 (3.0)	4.6 (2.3)
4;0–4;5	7.7 (3.0)	5.2 (2.3)
4;6–5;5	9.1 (2.5)	5.8 (2.1)
5;6–5;11	10.1 (1.4)	6.6 (1.6)
6;0–6;8	10.5 (0.8)	6.8 (1.4)

**Table 6. Raw score to percentile score conversion table for the sentence component of the test**

Raw score	Age group				
	3;4–3;11	4;0–4;5	4;6–5;5	5;6–5;11	6;0–6;8
0	<b>5</b>	<b>1</b>	<b>1</b>	<b>0</b>	<b>0</b>
1	<i>9</i>	<b>1</b>	<b>2</b>	<b>0</b>	<b>0</b>
2	<u>15</u>	<i>7</i>	<b>4</b>	<b>0</b>	<b>0</b>
3	<u>22</u>	<u>14</u>	<b>5</b>	<b>0</b>	<b>0</b>
4	31	<u>20</u>	<i>8</i>	<b>0</b>	<b>0</b>
5	45	27	<i>10</i>	<b>2</b>	<b>0</b>
6	55	32	<u>14</u>	<b>4</b>	<b>0</b>
7	64	36	<u>20</u>	<i>7</i>	<b>0</b>
8	73	50	27	<u>11</u>	<b>3</b>
9	85	61	37	21	<u>13</u>
10	94	81	60	44	40
11	100	100	100	100	100

Bold, <5%, italics, <10%; underline, <15%.

the child's performance relative to the standardization sample can be ascertained in terms of a percentile score.

*Children passing and/or failing the grammar and/or phonology subtests*

A final set of analyses was carried out on three different groups of children: those who were in our standardization sample; those who had been referred to the educational support services whose development in some aspects (language, cognitive, behaviour) was giving concern, and those who had already been identified as having SLI by SLTs and teachers in language resourced mainstream schools.

Tables 8–10 show the numbers (and percentages) of children in the standardization sample who passed or failed either or both subtests of the GAPS at each age group and as a proportion of the whole sample at the 5, 10 and 15% criterions.

**Table 7. Raw score to percentile score conversion table for the non-word component of the test**

Raw score	Age group				
	3;4–3;11	4;0–4;5	4;6–5;5	5;6–5;11	6;0–6;8
0	<b>5</b>	<b>4</b>	<b>4</b>	<b>0</b>	<b>0</b>
1	<u>12</u>	<i>8</i>	<i>6</i>	<b>1</b>	<b>1</b>
2	<u>19</u>	17	<i>9</i>	<b>2</b>	<b>1</b>
3	29	19	<u>13</u>	<b>5</b>	<b>3</b>
4	44	33	<u>22</u>	<u>11</u>	<i>7</i>
5	58	50	35	<u>22</u>	17
6	79	67	66	40	36
7	87	84	79	65	65
8	100	100	100	100	100

Bold, <5%, italics, <10%; underline, <15%.

**Table 8. Children who pass/failed either or both subtests at a 5% cut-off**

Age group			Non-words		
			Fail	Pass	Total
3;4–3;11	Sentences	fail	2 (2.3%)	2 (2.3%)	4 (4.6%)
		pass	2 (2.3%)	80 (93.1%)	82 (95.4%)
		total	4 (4.6%)	82 (95.4%)	86 (100.0%)
4;0–4;5	Sentences	fail	1 (0.9%)	0 (0.0%)	1 (0.9%)
		pass	3 (2.8%)	104 (96.3%)	107 (99.1%)
		total	4 (3.7%)	104 (96.3%)	108 (100.0%)
4;6–5;5	Sentences	fail	5 (1.7%)	11 (3.7%)	16 (5.4%)
		pass	6 (2.0%)	274 (92.6%)	280 (94.6%)
		total	11 (3.7%)	285 (96.3%)	296 (100.0%)
5;6–5;11	Sentences	fail	0 (0.0%)	4 (3.9%)	4 (3.9%)
		pass	5 (4.9%)	93 (91.2%)	98 (96.1%)
		total	5 (4.9%)	97 (95.1%)	102 (100.0%)
6;0–6;8	Sentences	fail	1 (1.3%)	1 (1.3%)	2 (2.6%)
		pass	1 (1.3%)	73 (96.1%)	74 (97.4%)
		total	2 (2.6%)	74 (97.4%)	76 (100.0%)
Overall	Sentences	fail	9 (1.3%)	18 (2.7%)	27 (4.0%)
		pass	17 (2.6%)	624 (93.4%)	641 (96.0%)
		total	26 (3.9%)	642 (96.1%)	668 (100.0%)

Tables 8–10 reveal that impairments in grammar and phonology for many children do not go hand in hand. For children falling into the lowest 10 or 15% on the GAPS subtests, the chances of having a phonological deficit, if you have a syntactic deficit, or vice versa is only around 0.5. Thus, of those children in the normal population who have some sort of language deficit, only half have both a phonological and a grammatical deficit. Surprisingly, the likelihood of having both a

**Table 9. Children who pass/failed either or both subtests at a 10% cut-off**

Age group			Non-words		
			Fail	Pass	Total
3;4–3;11	Sentences	fail	5 (5.8%)	3 (3.5%)	8 (9.3%)
		pass	2 (2.3%)	76 (88.4%)	78 (90.7%)
		total	7 (8.1%)	79 (91.9%)	86 (100.0%)
4;0–4;5	Sentences	fail	5 (4.6%)	2 (1.9%)	7 (6.5%)
		pass	4 (3.7%)	97 (89.8%)	101 (93.5%)
		total	9 (8.3%)	99 (91.7%)	108 (100.0%)
4;6–5;5	Sentences	fail	14 (4.7%)	14 (4.7%)	28 (9.4%)
		pass	13 (4.4%)	255 (86.2%)	268 (90.6%)
		total	27 (9.1%)	269 (90.9%)	296 (100.0%)
5;6–5;11	Sentences	fail	2 (2.0%)	5 (4.9%)	7 (6.9%)
		pass	6 (5.9%)	89 (87.2%)	95 (93.1%)
		total	8 (7.9%)	94 (92.1%)	102 (100.0%)
6;0–6;8	Sentences	fail	3 (4.0%)	4 (5.2%)	7 (9.2%)
		pass	2 (2.6%)	67 (88.2%)	69 (90.8%)
		total	5 (6.6%)	71 (93.4%)	76 (100.0%)
Overall	Sentences	fail	29 (4.4%)	28 (4.1%)	57 (8.5%)
		pass	27 (4.0%)	584 (87.5%)	611 (91.5%)
		total	56 (8.4%)	612 (91.6%)	668 (100.0%)

**Table 10. Children who pass/failed either or both subtests at a 15% cut-off**

Age group			Non-words		
			Fail	Pass	Total
3;4–3;11	Sentences	fail	7 (8.1%)	6 (7.0%)	13 (15.1%)
		pass	3 (3.5%)	70 (81.4%)	73 (84.9%)
		total	10 (11.6%)	76 (88.4%)	86 (100.0%)
4;0–4;5	Sentences	fail	8 (7.4%)	7 (6.5%)	15 (13.9%)
		pass	6 (5.6%)	87 (80.5%)	93 (86.1%)
		total	14 (13.0%)	94 (87.0%)	108 (100.0%)
4;6–5;5	Sentences	fail	22 (7.4%)	18 (6.1%)	40 (13.5%)
		pass	16 (5.4%)	240 (81.1%)	256 (86.5%)
		total	38 (12.8%)	258 (87.2%)	296 (100.0%)
5;6–5;11	Sentences	fail	3 (2.9%)	8 (7.9%)	11 (10.8%)
		pass	8 (7.9%)	83 (81.3%)	91 (89.2%)
		total	11 (10.8%)	91 (89.2%)	102 (100.0%)
6;0–6;8	Sentences	fail	4 (5.3%)	6 (7.9%)	10 (13.2%)
		pass	6 (7.9%)	60 (86.8%)	66 (86.8%)
		total	10 (13.2%)	66 (86.8%)	76 (100.0%)
Overall	Sentences	fail	44 (6.6%)	45 (6.7%)	89 (13.3%)
		pass	39 (5.8%)	540 (80.8%)	579 (86.6%)
		total	83 (12.4%)	585 (87.6%)	668 (100.0%)

grammar and phonological deficit in the most severely impaired group (the lowest 5%) reduces to around 0.33. This lack of concordance between grammar and phonological deficits accounts for why there are 6.5, 13.6 and 20% of the children in the sample at the 5, 10 and 15% criterion, respectively, where a level of concern over their language development is raised on this screening test.

Next we present the pass–fail results at 5%, 10% criterion (recommended level for specialist referral) and the 15% for the children referred to school support services (table 11) and those already identified as having some form of SLI (table 12).

Of these children (27 male and five female), who were referred to school support services as causing concern for a wide variety of reasons, around 60% failed the GAPS test at 15% criterion. All age groups are represented in this sample in low numbers. At 10% criterion, where referral to SLT services is recommended, there was still over 50% failure rate suggesting a substantial proportion of this population

**Table 11. Children referred to school support services who passed/failed one subtest or both**

Criterion			Non-words		
			Fail	Pass	Total
5%	Sentences	fail	6 (18.8%)	5 (15.6%)	11 (34.4%)
		pass	3 (9.3%)	18 (56.3%)	21 (65.6%)
		total	9 (28.1%)	23 (71.9%)	32 (100.0%)
10%	Sentences	fail	10 (31.2%)	7 (21.9%)	17 (53.1%)
		pass	1 (3.1%)	14 (43.8%)	15 (46.9%)
		total	11 (34.3%)	21 (65.7%)	32 (100.0%)
15%	Sentences	fail	10 (31.2%)	9 (28.2%)	19 (59.4%)
		pass	1 (3.1%)	12 (37.5%)	13 (40.6%)
		total	11 (34.3%)	21 (65.7%)	32 (100.0%)

**Table 12. Children with diagnosis of SLI who passed/failed one subtest or both**

Criterion	Non-words				
	Fail	Pass	Total		
5%	Sentences	fail	7 (41.2%)	3 (17.6%)	10 (58.8%)
		pass	0 (0.0%)	7 (41.2%)	7 (41.2%)
		total	7 (41.2%)	10 (58.8%)	17 (100.0%)
10%	Sentences	fail	7 (41.2%)	4 (23.5%)	11 (64.7%)
		pass	1 (5.9%)	5 (29.4%)	6 (35.3%)
		total	8 (47.1%)	9 (52.9%)	17 (100.0%)
15%	Sentences	fail	9 (52.9%)	3 (17.7%)	12 (70.6%)
		pass	0 (0.0%)	5 (29.4%)	5 (29.4%)
		total	9 (52.9%)	8 (47.1%)	17 (100.0%)

with poor attainment have some form of language and/or literacy difficulty as part of their problem which requires further diagnostic assessment. Of these children, 31% failed both subtests, whereas 22% failed only the grammatical subtest (at 10% cut-off). Only one child failed the non-word and passed the grammatical subtest.

From these children diagnosed as SLI by SLTs and other professionals, if they failed any test then they failed the grammar test and some children failed the phonology test too. One child failed only the phonology test and passed the sentences. The group comprised 12 males and five females with a predominantly even spread across the age groups. The five children with some form of diagnosed SLI who passed both subtests at the 10% criterion level were aged between 4;0 and 5;11 years. Had those passing been in the older age range we might have thought the test less sensitive at the older age group, but this does not appear to be the case. These results have not been statistically analysed further due to the small numbers. However, we discuss this small sample below.

## Discussion

The GAPS test has been designed to be a quick, simple screen of young children's language status. The screen aims to identify children in need of further assessment rather than to be definitive or diagnostic. Thus, for children whose scores fall in the lowest 10% when compared with their peers, further, detailed assessment by professionals is recommended. Retesting those children who scored at a borderline 15% cut-off should allow for any false-negative that might arise because of narrow scoring differentials. Indeed, with the older age group the precise 10–15 cut-off did not fall neatly within the scores and the level had to be set between the two at a raw score that differentiated the groups. We considered that the problems with small score differences were outweighed by the benefits of keeping the test short and easy to administer. In the older age range (6 years and over) TD children are predominantly grammatically competent on the test items and any error becomes significant. The repetition of non-words, however, shows a gradual development across the full age range of the test. The high correlations between the sentence repetition task and non-word repetition tasks and the possible increased sensitivity of a combined score will be explored in the follow-up.

The results of the standardization seemed to suggest that the rate of development between the ages of 4;6 and 5;5 years was less marked than at the other ages, especially for grammatical markers that are strongly associated with specific language impairment. One possible explanation is that this is due to other social and cognitive developments coming to the fore in this period of early educational experience, once complex grammar is in place. However, this interpretation requires further investigation to substantiate. An alternative explanation is that the items do not discriminate very precisely between these ages and that more items would help differentiate subtle changes occurring at these ages. Indeed morpho-syntactic development, even after 6;6 years continues with respect to complexity such as embedding and in more subtle and stylistic forms, after the initial core grammar is in place (van der Lely 2005). This possible weakness is balanced against our ethos of achieving a short and quick test.

The results of the pass–fail data reveal that the prevalence of grammatical and/or phonological deficits (i.e. those children in the lowest 5%) in the general population of children between 3;6 and 6;6 years is 6.5%. Tomblin and colleague's large epidemiological study (>7000 children) in the USA revealed a similar level of 7.4% prevalence of actual SLI in monolingual English-speaking kindergarten children (Tomblin *et al.* 1997), using a wider range of standardized, diagnostic tests. If we take the 10% criterion (the criterion we are recommending that children are referred for further assessment), then overall approximately 12.5% of children are identified as in need of further assessment. A total of 8.5% of these children were identified as having grammatical deficits with or without phonological deficits and 8.4% phonological deficits with half of these having no grammatical deficit. The prevalence of overall phonological deficits found in this standardization sample is close to the 10% incidence of dyslexia in the population (Snowling 2000). We cannot conclude that the children who fail on the GAPS phonological subtest are definitely those that go on to present with phonological developmental dyslexia. However, given our knowledge about dyslexia (e.g. Snowling 2000, Ramus 2001), such problems with phonological development are likely to put a child at risk for dyslexia. A similar pattern of prevalence of low performance in grammar and phonology but with raised incidence (19% overall) is found using the 15% criterion.

Although the test does not claim to be diagnostic, it does highlight children who are weaker in either morpho-syntax (grammar) or phonological processing or indeed both areas. Our findings provide further evidence for dissociation of language component impairment as discussed by van der Lely (2005) and Bishop *et al.* (2005). Such dissociations within language are important as each component deficit is likely to differentially impact on a child's education and would require qualitatively different remedial help. For example, a deficit in morpho-syntactic requires tailored intervention and could impact on understanding classroom instructions as well as reading comprehension, whereas an isolated phonological difficulty might impact more on decoding skills for reading and writing. Thus, it is noteworthy that there is generally a low level of co-morbidity between grammar and phonological deficits as revealed by our sentence and non-word subtests. This study revealed a co-morbidity at the 5% criterion of only 1.3% in the general population. This finding strongly concurs with Tomblin and colleagues large epidemiological study that revealed a similar low co-morbidity of 1.3% (Shriberg *et al.* 1999). The percentages of children in the general population identified with co-morbid weakness in grammar and phonology rises to 4.4% at the 10% criterion and 6.6% at the 15% criterion. In

other words, only 33% (5% criterion) to 50% (10 and 15% criterions) of children identified with either a grammar or phonological deficit also had a phonological or grammar deficit respectively. This finding is clearly surprising if auditory-phonological deficits are causing the grammatical deficits in young children from 3;6 years old as claimed by some theories of SLI (Tallal and Piercy 1973, Chiat 2001, Joanisse and Seidenberg 2003; cf. van der Lely 2005). However, the results are consistent with recent genetic evidence showing dissociation between genotypes for phonological and morpho-syntactic deficit (Bishop *et al.* 2005).

In the small sample of referred children who were exhibiting difficulties in the early school years (table 11), 56% were identified as having phonological and/or grammatical deficits at a level requiring referral. Therefore, there were proportionally more children in this group who were likely to be identified as having language/literacy difficulties than in the standardization sample (table 9). They did not have the same pattern of performance as the diagnosed SLI population (table 12) where 41% of children failed both subtests at the most stringent 5% cut-off as compared with 19% of the referred children. In this sample of children causing concern (table 11), only one child (3.1% of the small sample) failed just the non-word repetition test at 10% cut-off, showing a potentially very specific difficulty. The findings must be treated with caution due to the very small numbers, but suggest that perhaps only those children, in the early years of education, with the most severe and major deficits across a range of skills are likely to be referred to outside agencies. Children who are at significant risk of a more focused deficit such as specific reading difficulty due to phonological disability, in isolation from other language difficulties, might not cause too much concern at an early stage of literacy development where other reading strategies can compensate for poor phonological knowledge. Such children may also appear to be supported sufficiently by schools through a differentiated literacy curriculum at this stage.

In the small sample of children identified as having SLI, at the 10% cut-off (where referral would be made); 71% failed the phonological and/or grammatical subtests with 41% of these children failing both subtests, and 23.5% failing the grammatical subtest alone. Once again, there was only one child with an isolated phonological deficit who passed the grammatical items. Thus, the majority of children identified by SLTs as SLI appear to have both morpho-syntactic and phonological deficits. However, five (30%) of the 17 SLI children passed both subtests and therefore were not identified by the GAPS as being impaired on core grammatical or phonological abilities. This can be accounted for by the different forms of SLI, including those children with primarily lexical or pragmatic deficits which form the population in language resourced schools. Our percentages of children with phonological and/or grammatical deficits from this small sample concur with that found in previous research of SLI populations (Conti-Ramsden *et al.* 1997, Dockrell *et al.* 2001, Norbury and Bishop 2002). The children were also likely to have been recipients of good quality intervention in that they attended mainstream schools with specialist language resource provision. This may have accelerated progress in areas tapped by this simple screen.

Finally, these findings emphasize the value of our test in that it provides a first step in the process of identifying children with specific language difficulties and indeed those who might benefit from a differentiated literacy programme, much earlier than is now routine. Such early identification will mean additional, targeted



support can be given at a younger age, thereby hopefully helping prevent the development of wider educational problems.

This standardization report suggests that such a test can be successfully used by a wide range of people working with or caring for young children, not only those with professional training in language tests. The untrained testers, such as nursery nurses and teaching assistants, were slightly more lenient in their marking than the SLTs who used a wider range of scores and some of whom may have been slightly more discriminating in their scoring. Reports from the voluntary testers, both professional and non-professional were predominantly positive and they commented that the test was easy to use and they found the process a valuable exercise. A follow-up study will assess the accuracy of non-professionally trained testers through interrater reliability.

A follow-up of approximately 150 of those children involved in the original standardization testing is now underway. The children are being assessed on a variety of standardized language and simple reading tests as well as being retested on the GAPS. This study will therefore track the development of children who both passed and failed the GAPS and confirm whether the former group are continuing to progress well and whether the group that failed continue to have language or literacy difficulties.

### **Conclusions and implications**

The elicited imitation of sentences that contain grammatical markers of SLI and non-words that vary in phonological complexity provides the basis for this standardized GAPS test. The simple tick/cross scoring method has been combined with a reliable standardization. The prevalence of grammatical and phonological problems in the general population identified by this quick screening test concurs closely with previous large epidemiological studies based on a lengthy battery of language tests. Thus, the GAPS should form a valuable addition to the battery of full language tests available, for the most part because it is short, simple and does not necessarily require specialist training to administer. However, further training if necessary could be provided to many support workers within the prevailing system with little added cost. The focus within education and health, and the large government funded 'Sure-Start scheme' is towards more collaborative practice and sharing of expertise; this screen therefore fits well with this ethos.

### **What this paper adds**

#### *What is already known about this subject*

Approximately 7% of children have language impairment, and 10% suffer from reading or writing impairment which significantly affects such children reaching their potential. Current standardized tests, requiring professional administration, often lasting more than 30 min can identify such children. However there is no short, (quick) standardized screening assessment, using key grammatical markers of language impairment and phonological markers for children at risk for dyslexia in the preschool and early school years that could be routinely administered by a concerned parent or professional alike.

### **What this study adds**

This study provides details of a new standardised, quick to administer test for professionals and non-professionals: the Grammar and Phonology Screening (GAPS) test identifies children at risk of language and literacy deficits in similar proportions to those found using much lengthier procedures.

### **Acknowledgements**

The GAPS test could not have come to fruition without the wonderful support throughout the UK the authors had had from schools, parents, students, and speech and language therapists, as well as from all the children who helped in many ways to undertake the standardization. In addition, particular thanks are due to Carmen Herzog, Potsdam University, who during her 2-month visit to the DLDCN Centre carried out a large part of the testing for the pilot study and to Peter Hudspith, artist. The authors are also very grateful to the De Laszlo Foundation for its support in the latter stages of the project. H. v. d. L. was supported by a Wellcome Trust University Award (Ref. 063713), which is gratefully acknowledged.

### **References**

- BISHOP, D. V. M., 1997, *Uncommon Understanding: Comprehension in Specific Language Impairment* (Hove: Psychology Press).
- BISHOP, D. V., ADAMS, C. V. and NORBURY, C. F., 2004, Using nonword repetition to distinguish genetic and environmental influences on early literacy development: a study of 6-year-old twins. *American Journal of Medical Genetics*, **129B**, 94–96.
- BISHOP, D. V. M., ADAMS, C. V. and NORBURY, C. F., 2005, Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain and Behavior* (in press).
- BISHOP, D., BRIGHT, P., JAMES, C., BISHOP, S. and VAN DER LEY, H., 2000, Grammatical SLI: a distinct subtype of developmental language disorder? *Applied Psycholinguistics*, **21**, 159–181.
- BISHOP, D. V. M., NORTH, T. and DONLAN, C., 1996, Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *Journal of Child Psychology and Psychiatry*, **37**, 391–403.
- BROWN, R., 1957, Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology*, **55**, 1–5.
- CHIAT, S., 2001, Mapping theories of developmental language impairment: premises predictions and evidence. *Language and Cognitive Processes*, **16**, 113–142.
- CHOMSKY, N., 1986, *Knowledge of Language: Its Nature, Origin and Use* (New York, NY: Praeger).
- CHOMSKY, N., 1995, *The Minimalist Program* (Cambridge MA: MIT Press).
- CHOMSKY, N. and HALLE, M., 1968, *The sound pattern of English* (New York: Harper and Row. Boston: MIT Press).
- CLAHSEN, H. and ALMAZAN, M., 1998, Syntax and morphology in Williams Syndrome. *Cognition*, **68**, 167–198.
- CLAHSEN, H., BARTKE, S. and GOELLNER, S., 1997, Formal features in impaired grammars: a comparison of English and German SLI children. *Journal of Neurolinguistics*, **10**, 151–171.
- CONTI-RAMSDEN, G., 2003, Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech–Language Hearing Research*, **46**, 1029–1037.
- CONTI-RAMSDEN, G. and HESKETH, A., 2003, Risk markers for SLI: a study of young language-learning children. *International Journal of Language and Communication Disorders*, **38**, 251–263.
- CONTI-RAMSDEN, G., BOTTING, N. and FARAGHER, B., 2001, Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, **42**, 741–748.
- CONTI-RAMSDEN, G., CRUTCHLEY, A. and BOTTING, N., 1997, The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research*, **40**, 765–777.

- CRAIN, S. and THORNTON, R., 1998, *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics* (Cambridge, MA: MIT Press).
- DALE, P. S., PRICE, T. S., BISHOP, D. V. and PLOMIN, R., 2003, Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech–Language Hearing Research*, **46**, 544–560.
- DEPARTMENT OF EDUCATION AND SKILLS, 2001, *The National Literacy Strategy*. Available online at <http://www.standards.dfes.gov.uk/literacy/publications/research/63541/>
- DOCKRELL, J., MESSER, D. and GEORGE, R., 2001, Patterns of naming objects and actions in children with word finding difficulties. *Language and Cognitive Processes*, **16**, 261–286.
- DUNN, L., DUNN, L., WHETTON, C. and BURLEY, J., 1997, The British Picture Vocabulary Scale, 2nd edn. Unpublished manuscript (Windsor).
- EBBELS, S., 2005, Argument structure in specific language impairment: from theory to therapy. Unpublished PhD, University College London.
- EBBELS, S. and VAN DER LELY, H. K. J., 2001, Meta-syntactic therapy for children with severe persistent SLI using visual coding. *International Journal of Language and Communication Disorders*, **36**, 345–350.
- EDWARDS, S., FLETCHER, P., GARMAN, M., HUGHES, A., LETTIS, C. and SINKA, I., 1997, *The Reynell Developmental Language Scales III* (Windsor: NFER-Nelson).
- FEY, M. E., CLEAVE, P. L. and LONG, S. H., 1997, Two models of grammar facilitation in children with language impairments: phase. *Journal of Speech Language and Hearing Research*, **2**, 5–19.
- FRIEDMANN, N. and NOVOGRODSKY, R., 2004, The acquisition of Relative clause comprehension in Hebrew: a study of SLI and normal development. *Journal of Child Language*, **31**, 661–681.
- GALLON, N., HARRIS, J. and VAN DER LELY, H. K. J., 2005, An investigation into phonological complexity using non-word repetition tasks in children with G-SLI (submitted).
- GATHERCOLE, S. and ADAMS, A., 1993, Phonological working memory in very young children. *Developmental Psychology*, **29**, 770–778.
- GATHERCOLE, S. and BADDELEY, A., 1990, Phonological memory deficits in language disordered children: is there a causal connection? *Journal of Memory and Language*, **29**, 336–360.
- GATHERCOLE, S. and BADDELEY, A., 1996, *The Children's Test of Non-Word Repetition* (London: Psychological Corporation).
- HAMANN, C., PENNER, Z. and LINDNER, K., 1998, German impaired grammar: the clause structure revisited. *Language Acquisition*, **7**, 193–245.
- HARRIS, J., 1994, *English Sound Structure* (Oxford: Blackwell).
- HAUSER, M. D., CHOMSKY, N. and FITCH, W. T., 2002, The faculty of language: what is it, who has it, and how did it evolve? *Science*, **298**, 1569–1579.
- HAYNES, C. and NAIDOO, S., 1991, *Children with Specific Speech and Language Impairment* (Oxford: MacKeith).
- HEATH, S. and HOGGEN, J. H., 2004, Cost-effective prediction of reading difficulties. *Journal of Speech and Hearing Research*, **47**, 751–765.
- JOANISSE, M. F. and SEIDENBERG, M. S., 2003, Phonology and syntax in specific language impairment: evidence from a connectionist model. *Brain and Language*, **86**, 40–56.
- LAING, G., LAW, J., LEVIN, A. and LOGAN, S., 2002, Evaluation of a structured test and a parent led method for screening for speech and language problems: prospective population based study. *British Medical Journal*, **325**, 1152–1155.
- LAW, J., GARRETT, Z. and NYE, N., 2003, Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane review*, no. 1.
- LAW, G. and BISHOP, D. V., 2003, A comparison of language abilities in adolescents with Down's syndrome and children with specific language impairment. *Journal of Speech–Language Hearing Research*, **46**, 1324–1339.
- LEONARD, L., 1998, *Children with Specific Language Impairment* (Cambridge, MA: MIT Press).
- LUST, B., FLYNN, S. and FOLEY, C., 1996, What children know about what they say: elicited imitation as a research method for assessing children's syntax. In D. McDaniel, C. McKlee and H. Smith Cairns (eds), *Methods for Assessing Children's Syntax* (Cambridge, M: MIT Press), pp. 55–76.
- MARSHALL, C. R., EBBELS, S., HARRIS, J. and VAN DER LELY, H. K. J., 2002, Investigating the impact of prosodic complexity on the speech of children with Specific Language Impairment. In R. Vermeulen and A. Neeleman (eds), *UCL Working Papers in Linguistics*, Vol. **14**, pp. 43–68.
- MARSHALL, C., HARRIS, J. and VAN DER LELY, H. K., 2003, The nature of phonological representations in children with Grammatical-Specific Language Impairment (G-SLI). In D. Hall, T. Markopoulos, A. Salamoura and S. Skoufaki (eds), *The University of Cambridge First Postgraduate*

*GAPS test to assess key speech markers in young children*

- Conference in Language Research* (Cambridge: Cambridge Institute of Language Research, University of Cambridge), pp. 511–517.
- MARSHALL, C. and VAN DER LELY, H., A challenge to current models of past tense inflection: the impact of phonotactics. *Cognition* (in press).
- McDANIEL, D., McKEE, C. and SMITH CAIRNS, H., 1996, *Methods for Assessing Children's Syntax* (Cambridge, MA: MIT Press).
- MORGAN, J. L. and DEMUTH, K., 1996, *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (Mahwah, NJ: Lawrence Erlbaum).
- NATION, K., CLARKE, P., MARSHALL, C. M. and DURAND, M., 2004, Hidden language impairments in children: parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language and Hearing Research*, **47**, 199–211.
- NATIONAL INSTITUTE FOR LITERACY, 2001, *Put Reading First: The Research Building Blocks for Teaching Children to Read*, NIFL, National Institute of child Health and Human Development (NICHD) and US Department of Education (available at [http://www.nifl.gov/partnershipforreading/publications/reading\\_first1.html](http://www.nifl.gov/partnershipforreading/publications/reading_first1.html)).
- NÄÄTÄNEN, R., LEHTOKOSKI, A., LENNES, M., CHEOUR, M., HUOTILAINEN, M., IIVONEN, A., VAINIO, M., ALKU, P., ILMONIEMI, R. J., LUUK, A., ALLIK, J., SINKKONEN, J. and ALHO, K., 1997, Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, **385**, 432–434.
- NELSON, K. E., CAMARATA, S. M., WELSH, J., BUTKOVSKY, L. and CAMARATA, M., 1996, Effects of imitative and conversational recasting treatment on the acquisition of grammar in children with specific language impairment and younger language-normal children. *Journal of Speech and Hearing Research*, **39**, 850–859.
- NEWMAN, S., 1996, Working on both sides of the fence: the effect of dual qualification on collaborative working practice. *Child Language Teaching and Therapy*, **2**, 39–47.
- NORBURY, C. and BISHOP, D., 2002, Inferential processing and story recall in children with communication problems: a comparison of specific language impairment, pragmatic language impairment and high-functioning autism. *International Journal of Language and Communication Disorders*, **37**, 227–251.
- NORBURY, C., BISHOP, D. and BRISCOE, J., 2001, Production of English finite verb morphology: a comparison of SLI and mild-to-moderate hearing impairment. *Journal of Speech, Language and Hearing Research*, **44**, 165–178.
- NORBURY, C., BISHOP, D. and BRISCOE, J., 2002, Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics*, **23**, 247–268.
- O'HARA, M. and JOHNSTON, J., 1997, Syntactic bootstrapping in children with specific language impairment. *European Journal of Disorders of Communication*, **2**, 189–205.
- OFFICE OF NATIONAL STATISTICS, 2000a, *Standard Occupational Classification*. Vol. 1: *Structure and Descriptions of Unit Groups* (available at: [http://www.statistics.gov.uk/methods\\_quality/ns\\_sec/downloads/SOC2000\\_Vol1\\_V5.pdf](http://www.statistics.gov.uk/methods_quality/ns_sec/downloads/SOC2000_Vol1_V5.pdf)).
- OFFICE OF NATIONAL STATISTICS, 2000b, *Standard Occupational Classification*. Vol. 2: *The Coding Index* (available at: [http://www.statistics.gov.uk/methods\\_quality/ns\\_sec/downloads/SOC2000\\_Vol2\\_V6.pdf](http://www.statistics.gov.uk/methods_quality/ns_sec/downloads/SOC2000_Vol2_V6.pdf)).
- OFFICE OF NATIONAL STATISTICS, 2005, *National Statistics*. (available at: <http://www.statistics.gov.uk/census2001/>).
- PICKSTONE, C., HANNON, P. and FOX, L., 2002, Surveying and screening preschool language development in community focussed intervention programmes: a review of instruments. *Child Care, Health and Development*, **28**, 251–264.
- PINKER, S., 1994, *The Language Instinct* (New York, NY: W. Morrow).
- PINKER, S., 1999, *Words and Rules: The Ingredients of Language* (London: Weidenfeld & Nicolson).
- PRECIOUS, A. and CONTI-RAMSDEN, G., 1988, Language-impaired children's comprehension of active versus passive sentences. *British Journal of Disorders in Communication*, **23**, 229–243.
- QUALITY CURRICULUM AUTHORITY, 2003, The speaking and listening curriculum (available at <http://www.standards.dfes.gov.uk/literacy/publications/framework/818497/>).
- RAMUS, F., 2001, Outstanding questions about phonological processing in dyslexia. *Dyslexia*, **7**, 197–216.
- RAMUS, F., 2003, Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, **13**, 212–218.
- RICE, M., 2004, Language growth of children with SLI and unaffected children: timing mechanisms and linguistic distinctions. In A. Brugos, L. Micciulla and C. Smith (eds), *Proceeding of the 28th Annual Boston University Conference on Language Development* (Somerville, MA: Cascadilla), **vol. 1**, 28–49.

- RICE, M. and WEXLER, K., 1996, Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, **39**, 1239–1257.
- RING, M. and CLAHSEN, H., 2003, Distinct patterns of language impairment in Down Syndrome, Williams Syndrome, and SLI: the case of syntactic chains. Unpublished manuscript, University of Essex, Colchester.
- ROY, P. and CHIAT, S., 2004, A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: evidence from typically developing children. *Journal of Speech–Language Hearing Research*, **47**, 223–234.
- SHRIBERG, L. D., TOMBLIN, J. B. and MCSWEENEY, J. L., 1999, Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech–Language Hearing Research*, **42**, 1461–1481.
- SNOWLING, M., 2000, *Dyslexia*. 2nd edn (Oxford: Blackwell).
- STAVRAKAKI, S., 2001, Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language*, **77**, 419–431.
- STAVRAKAKI, S., 2002, A-bar movement constructions in Greek children with SLI: evidence for deficits in the syntactic component of language. In E. Fava (ed.), *Clinical Linguistics: Theory and Applications in Speech Pathology and Therapy. Current Issues in Linguistic Theory* (Amsterdam: John Benjamins), pp. 131–153.
- TALLAL, P. and PIERCY, M., 1973, Developmental aphasia: impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, **11**, 389–398.
- THORNTON, R., 1995, Referentiality and Wh-movement in child English: Juvenile D-Linkuency. *Language Acquisition*, **4**, 139–175.
- TOMBLIN, J. B., RECORDS, N. L., BUCKWALTER, P., ZHANG, X., SMITH, E. and O'BRIEN, M., 1997, Prevalence of specific language impairment in kindergarten children. *Journal of Speech–Language Hearing Research*, **40**, 1245–1260.
- VAN DER LELY, H. K. J., 1996a, Specifically language impaired and normally developing children: verbal passive versus adjectival passive sentence interpretation. *Lingua*, **98**, 243–272.
- VAN DER LELY, H. K. J., 1996b, *The Test of Active and Passive Sentences (TAPS)* (available from author: Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London).
- VAN DER LELY, H. K. J., 1997, *Advanced-Syntactic test of Pronominal reference (A-STOP)* (available from author: Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London, UK).
- VAN DER LELY, H. K. J., 1998, SLI in children: movement, economy and deficits in the computational syntactic system. *Language Acquisition*, **72**, 161–192.
- VAN DER LELY, H. K. J., 1999, *Verb Agreement and Tense Test (VATT)*, London (available from author: Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London, UK).
- VAN DER LELY, H. K. J., 2004, Evidence for and implications of a domain-specific grammatical deficit. In L. Jenkins (ed.), *The Genetics of Language* (Oxford: Elsevier), pp. 117–144.
- VAN DER LELY, H. K. J., 2005, Domain-specific cognitive systems: insight from grammatical-specific language impairment. *Trends in Cognitive Sciences*, **9**, 53–59.
- VAN DER LELY, H. K. J. and BATTILL, J., 2003, Wh-movement in children with grammatical SLI: a test of the RDDR hypothesis. *Language*, **79**, 153–181.
- VAN DER LELY, H. K. J. and HARRIS, J., 1999, *The Test of Phonological Structure (TOPhS)* (London: available from author: Centre for Developmental Language Disorders & Cognitive Neuroscience, University College London).
- VAN DER LELY, H. K. J., ROSEN, S. and ADLARD, A., 2004, Grammatical language impairment and the specificity of cognitive domains: relations between auditory and language abilities. *Cognition*, **94**, 167–183.
- VAN DER LELY, H. K. J., ROSEN, S. and MCCLELLAND, A., 1998, Evidence for a grammar-specific deficit in children. *Current Biology*, **8**, 1253–1258.
- VAN DER LELY, H. K. J. and STOLLWERK, L., 1997, Binding theory and grammatical specific language impairment in children. *Cognition*, **62**, 245–290.
- WIG, E. H., SECORD, W. and SEMEL, E., 2000, *Clinical Evaluation of Language Fundamentals — Preschool* (London: Psychological Corporation).