

Grammatical SLI: A distinct subtype of developmental language impairment?

D. V. M. BISHOP
University of Oxford

P. BRIGHT, C. JAMES, and S. J. BISHOP
MRC Cognition and Brain Sciences Unit

H. K. J. VAN DER LELY
Birkbeck College

ADDRESS FOR CORRESPONDENCE

D. V. M. Bishop, University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK. Email: dorothy.bishop@psy.ox.ac.uk

ABSTRACT

Grammatical specific language impairment (G-SLI) has been proposed as a distinct subtype of language impairment. We assessed a large sample of twins between the ages of 7 and 13 years on language comprehension tests sensitive to G-SLI. The sample included 37 same-sex twin pairs selected for the presence of language impairment (LI) in one or both twins and 104 twin pairs from the general population. The number and patterns of errors in those with LI replicated findings from previous studies of G-SLI. Qualitative markers of G-SLI were derived from the tests. Out of 144 children for whom complete data were available, 2 scored positive on all five markers and 9 scored positive on four of the five markers. Most children who made grammatical errors characteristic of G-SLI had co-occurring deficits in other areas. These results raise questions for theories that treat this disorder as an all-or-none modular deficit.

Deficits in the use and understanding of grammar are a hallmark of specific language impairment (SLI). Most research on grammatical problems has concentrated on expressive difficulties with grammatical morphology (see Leonard, 1998). However, in at least a subset of children, the ability to understand the meaning signaled by grammatical devices such as inflectional endings or word order is also impaired (Adams, 1990; Bishop, 1979, 1982). In a series of papers, Van der Lely and colleagues (Van der Lely, 1994, 1996; Van der Lely & Stollwerck, 1997) have argued that there is a homogeneous subtype of grammatical SLI (G-SLI), which is characterized by a “representational deficit for dependent relationships” (RDDR) in the computational syntactic system. This leads to inconsistency in forming or using syntactic structures that involve dependent rela-

tionships between syntactic elements, so that an affected child will rely on other, nonsyntactic cues to interpret meaning in sentences. It is argued that, on comprehension tests that require computation of dependent relationships, children with G-SLI make distinctive errors, which cannot simply be attributed to immature language development. Another claim made by Van der Lely and Stollwerck (1996) is that the etiology of G-SLI is strongly genetic, possibly controlled by an autosomal dominant gene. This is not to imply that all grammar is controlled by a single gene, but rather that a genetic deficit can have an impact on specific mechanisms or representations underlying grammar.

The RDDR hypothesis

The RDDR hypothesis maintains that there is a subset of children with SLI who have difficulty in using and understanding sentence elements that mark syntactic dependencies. Examples of such dependencies in English are subject–verb agreement (e.g., “I go” but “he goes”; “he is running” vs. “they are running”) and marking of pronoun case (e.g., the contrast between *he* vs. *him*), which depends on the syntactic relationship between a noun phrase and the verb phrase. Both of these are aspects of grammar where expressive errors are commonly seen in children with SLI.

The RDDR hypothesis has evolved over time, while retaining the core notion that the underlying deficit in G-SLI is confined to the linguistic computational system rather than involving a more general processing capacity limitation. In the 1997 version of the hypothesis, Van der Lely and Stollwerck argued that a child with G-SLI generates a syntactic representation for an incoming sentence that does not fully specify the relationship between sentence constituents. Elementary local dependencies may be built, but complex long-distance dependencies are not. The result is that the representation may be ambiguous and open to more than one interpretation. If that were the whole explanation, then one might expect children with G-SLI to perform at chance when given complex sentences where long-distance syntactic dependencies determine the meaning. In fact, performance is typically above chance, even when there are no semantic or pragmatic cues to meaning. More recently, Van der Lely (in press) has proposed a modified RDDR hypothesis in which comprehension errors reflect the failure to apply the grammatical operation of movement in contexts when it is obligatory. Instead, children with G-SLI treat movement as optional. The net result is that sentences with complex syntactic dependencies will sometimes be interpreted correctly and sometimes incorrectly.

In the current article, we focus on two aspects of comprehension that have been used by Van der Lely as evidence for the RDDR hypothesis. The first is the interpretation of active and passive sentences. In reversible sentences (e.g., “the boy chases the dog”), either noun phrase could plausibly take the theme of agent (i.e., boys can chase dogs and dogs can chase boys). Correct interpretation requires syntactic analysis to determine who does what to whom, which would not be the case if a nonreversible sentence (e.g., “the man eats the apple”) were used. Both normally developing and language-impaired children find reversible passive sentences harder to decode than reversible active sentences. Some chil-

dren with SLI will consistently misinterpret reversible passive sentences by assigning the thematic role “agent” to the first-named entity (e.g., “the boy is chased by the dog” is interpreted as “boy chase dog”) (Bishop, 1982). More commonly, however, a mixture of correct and reversed interpretations is given to the reversible passive sentences. Van der Lely and Harris (1990) demonstrated this using both multiple-choice picture selection and an acting-out task to assess comprehension. In the latter task, the child was given small toys and asked to make them perform a described action (e.g., “the bear is hit by the tiger”). Van der Lely and Harris concluded that the processes of performing a syntactic analysis and relating this to thematic roles posed particular problems for children with G-SLI.

In an unpublished thesis, Van der Lely (1990) noted that children with G-SLI performed better at acting out short passives (e.g., “the boy is chased”) than full passives (e.g., “the boy is chased by the girl”). At first glance, this seems incompatible with an underlying deficit in representing dependent relations, because the short passives are regarded as having a complex underlying structure with a noncanonical relation between thematic roles and grammatical functions. However, Van der Lely (1994) suggested that children with G-SLI could succeed with these structures if, instead of treating the verb as causative, they interpreted it as a stative verb, like an adjective. For instance, on hearing “the toy is broken,” one can regard “broken” as a state/adjective, so that the sentence has the same type of structure as “the toy is red” or “the toy is big.” Indeed, a word like “broken” can play an explicitly adjectival role in sentences such as “the broken toy is in the cupboard.” This adjectival passive interpretation is syntactically much simpler, as only one thematic role needs to be assigned. In contrast, the verbal passive has two thematic roles to assign. In a study using picture arrays such as the one shown in Figure 1, Van der Lely (1996) obtained evidence that children with G-SLI tend to adopt an adjectival interpretation of short ambiguous passives.

Another aspect of syntax investigated by Van der Lely and Stollwerck (1997) is Binding Principles, which determine pronominal reference. For instance, in “Baloo Bear says Mowgli is tickling him,” “him” refers to Baloo Bear; however, in “Baloo Bear says Mowgli is tickling himself,” the reflexive “himself” refers to Mowgli. These pronominal forms occupy exactly the same sentence position; their interpretation must depend on something other than just word order. According to Principles and Parameter theory (Chomsky, 1981), Binding Principles specify that a reflexive must be coreferential with a “local” antecedent (broadly speaking, this means it will refer to a noun phrase within the same clause). Thus, the antecedent must have a particular syntactic relationship to the reflexive. In contrast, a personal pronoun cannot refer to a noun phrase in a local syntactic relationship. A child who failed to compute syntactic relationships accurately would have difficulty in applying these Binding Principles. Van der Lely and Stollwerck (1997) demonstrated that this was the case for children with G-SLI, who were tested in a judgment paradigm in which they had to respond “yes” or “no” depending on whether a picture, such as those in Figure 2, matched a spoken sentence. These children were able to use nonsyntactic cues to meaning, such as semantic gender of the pronoun (e.g., him = male), and so they had little

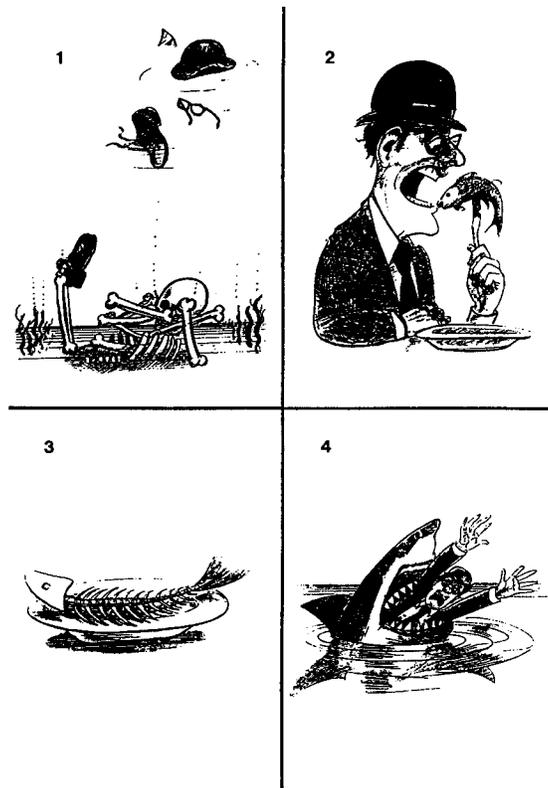


Figure 1. Sample item from the Test of Active and Passive Sentences: "The fish is eaten by the man."

difficulty in decoding analogous sentences where the two named participants were different genders. However, in "Baloo Bear says Mowgli is tickling him," on a substantial number of occasions they responded "yes" even when Mowgli was depicted tickling himself.

This study of intuitive knowledge of Binding Principles also involved a comparison between children with G-SLI and younger, normally developing children. In reviewing this work, Bishop (1997b) noted that, although 11-year-old children with G-SLI once again did more poorly than language ability matched controls, their pattern of performance did resemble that of normally developing children below 6 years of age, as previously reported. However, normally developing children who made these types of errors also tended to make errors on other sentences that did not demand knowledge of Binding Principles. In contrast, the children with G-SLI showed a discontinuity between their relatively good performance in comprehending control sentences and their poor performance on sentences that tested knowledge of Binding Principles. Thus, the pattern of performance on sentences requiring grammatical versus nongrammatical

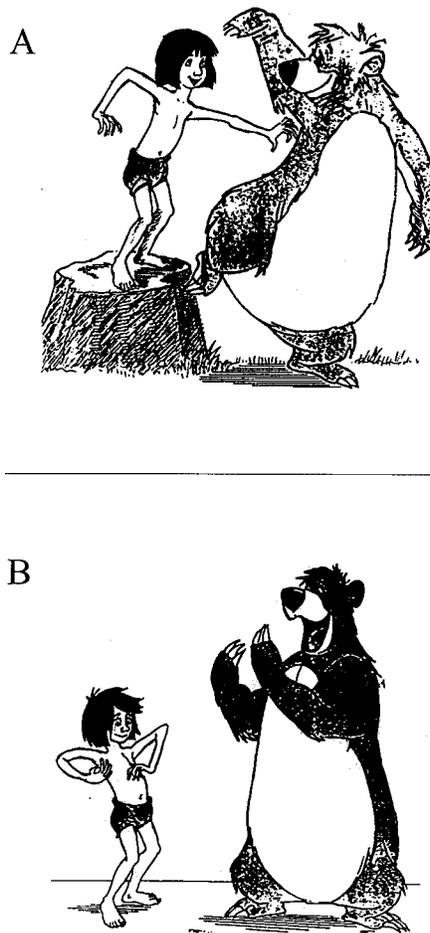


Figure 2. Two sample items from the Advanced Syntactic Test of Pronominal Reference. Each item consists of a single picture and spoken sentence; the child's task is to say "yes" if the picture and sentence match and "no" if they do not.

language knowledge by children with G-SLI did not resemble that of normally developing children at any age.

In 1997, Van der Lely gave a detailed description of G-SLI and reported a case study of a boy, AZ, who exemplified this subtype of SLI. AZ had a history of extreme language delay, and at 5 years of age his expressive vocabulary consisted only of "Mummy," "Daddy," and "Gangan" (Granny). There was nothing remarkable about his early perinatal or medical history. Both his father and his paternal uncle had evidence of residual language difficulties. AZ was studied between the ages of 10 and 12 years at the special school for children with speech and language impairments that he had attended since the age of 7

years. On standardized nonverbal tasks, AZ scored in the top 10% or above in relation to chronological age; his performance on Raven's Matrices, a nonverbal reasoning test, was at the 94th centile for his age. He had no obvious phonological impairment in expressive speech. On standardized tests, his morphosyntactic abilities were similar to those of control children just under 6 years, whereas his vocabulary was similar to control children just under 7 years. He demonstrated major difficulties on a task designed to elicit inflectional morphology and on a grammaticality judgment task involving verb inflections. His utterances were restricted to syntactically simple sentences. He did poorly on the test of knowledge of Binding Principles, except where he could use nonsyntactic cues such as gender to guide performance. In striking contrast to these impaired language skills, AZ had no difficulty with a test of pragmatic awareness, which involved detecting utterances that violated conversational maxims. He also performed at age level or better on a test of verbal reasoning and on a nonverbal test of transitive inference. When his auditory abilities were tested at the age of 15 years, he showed essentially normal performance on tests of temporal order judgment and backward masking (Rosen, Van der Lely, & Dry, 1997). More recently, Van der Lely, Rosen, and McClelland (1998) gave a brief account of six additional cases of G-SLI.

G-SLI as a specific subgroup with a genetic etiology

Van der Lely and Stollwerck (1996) analyzed family pedigree data from children who met criteria for G-SLI, all of whom were attending a special residential school for children with SLI. The pattern of familiarity of affected individuals was compatible with a genetic etiology for this disorder, with an autosomal dominant mode of inheritance. Twin data would provide a stronger test of this hypothesis, because they allow one to separate the effects of shared environment and shared genes in related individuals. For polymorphic genes (i.e., those that show individual variation within the species), monozygotic (MZ) twins are genetically identical, whereas dizygotic (DZ) twins share on average 50% of their genes, just like any other pair of siblings. For twins growing up together in the same household, there will be many shared environmental experiences, as well as a shared prenatal environment in the womb. These environmental factors will account for some similarities between twin pairs. However, insofar as genes are important in causing a trait or disorder, we would expect MZ twins to show closer similarity to one another than DZ twins. According to Van der Lely and Stollwerck's (1996) account, G-SLI is caused by one or more genes and is an all-or-none condition. Thus, if we classify children in terms of whether they show the pattern of comprehension deficits that characterizes G-SLI, we would expect to find higher concordance for MZ than for DZ twins.

Other researchers have presented evidence for a strong genetic contribution to SLI (e.g., Bishop, North, & Donlan, 1995; Lewis & Thompson, 1992; Tomblin & Buckwalter, 1994), but they have not attempted to distinguish a grammatical subtype of SLI. There is particular interest in the question of whether recognition of G-SLI as a distinct phenotype will clarify our understanding of the genetic basis of children's language disorders.

AIMS OF THE CURRENT STUDY

The general goal of the current study was to assess the validity of G-SLI as a grammatical subtype within the SLI population, characterized by comprehension difficulties as predicted by the RDDR hypothesis. More specifically, our aims were as follows:

1. to replicate the studies of Van der Lely (1996) and Van der Lely and Stollwerck (1997), which demonstrated that there is a subset of children with SLI who make the specific comprehension errors predicted by the RDDR;
2. to consider whether such children show a selective deficit in grammar, similar to that described in AZ, or whether their deficits co-occur with other difficulties, either in the nonsyntactic aspects of language or on nonverbal tasks; and
3. to use classic twin study methodology to assess the heritability of a deficit in grammatical comprehension, as measured by Van der Lely's tests.

The group of twins participating in the current study was recruited from two samples: in sample A children were selected on the basis that at least one member of the twin pair had a language impairment; sample B was a general population sample of twins. By including both samples, we were able to look at error patterns in comprehension in a large sample of normally developing children as well as in a group of children with SLI.

PARTICIPANTS

Full details of the children taking part in this study are reported by Bishop et al. (1999). In brief, we included 37 same-sex twin pairs selected for the presence of language impairment in one or both twins (sample A) and an unselected sample of 104 twin pairs from the general population (sample B). All 282 children were between the ages of 7 to 13 years. Twins from sample A had originally been recruited and assessed in 1989–1991; the data reported here are from a reassessment carried out in 1995. Twins from sample B were recruited from state schools in Hertfordshire and Cambridgeshire and were assessed during 1995–1996.

Although many of the analyses reported here used data from all children from both samples, preliminary analyses focused only on those who were unambiguously language-impaired (LI) or language-normal (LN). Table 1 summarizes the criteria used for each category and the characteristics of those who did not meet the criteria for either LI or LN. In categorizing children from sample A, the question arose as to how to handle a child who met the criteria for LI when first assessed but no longer did so when reassessed for this study. We decided to include such children in the LI category for this study. This meant that we included in the LI group some children who had demonstrated normal nonverbal ability on the initial assessment but who fell below cutoff when reassessed.

Twin zygosity was determined using a standard questionnaire that relies on similarities in physical characteristics (Nichols & Bilbro, 1966). Genetic fingerprinting from blood samples or cheek scrapings were used in cases where the questionnaire was inconclusive.

Table 1. *Classification of children in terms of language and nonverbal status*

Classification	Definition	<i>N</i>	<i>M (SD)</i> language ^a	<i>M (SD)</i> nonverbal ^b
Language-impaired (LI)	<ul style="list-style-type: none"> • nonverbal scaled score above 80 on Raven's Matrices and short form PIQ • scaled scores below 80 on two or more core language tests, including at least one receptive language test^c 	54	77.9 (7.66)	90.2 (9.91)
Language-normal (LN)	<ul style="list-style-type: none"> • nonverbal scaled score between 80 and 120 on Raven's Matrices and short form PIQ • all core language test scaled scores over 80 • no history of speech therapy, parental concern about language, or literacy problems 	76	103.7 (7.67)	101.4 (8.03)
Not meeting the criteria for LI or LN	Marginal: expressive language impairment only or score below 80 on a single language test	12	87.6 (5.06)	101.9 (8.30)
	Resolved: history of speech-language therapy, parental concern about language, or remedial reading, but normal language test scores	73	96.5 (8.47)	102.2 (11.45)
	High IQ: Nonverbal scaled score above 120 on Raven's Matrices or WISC-III short form PIQ	35	109.6 (7.37)	118.3 (5.99)
	Low IQ: nonverbal scaled score below 80 on Raven's Matrices of WISC-III short form PIQ	32	88.9 (9.90)	79.7 (5.93)

^aMean score on four core language tests.

^bMean score on Raven's Matrices and short form PIQ.

^cChildren in sample A had been assessed on two occasions, and this category was used if the criterion was met on either assessment.

LANGUAGE AND COGNITIVE ASSESSMENT

The individual assessment of each child lasted between 75 and 90 minutes and was carried out at home or in a quiet room at school. The battery of tests included some experimental measures that will not be reported here as well as four core language tests, which were used to classify children's language. Also included were two tests devised by Van der Lely and colleagues: the Test of

Active and Passive Sentences (TAPS) and a shortened version of the Advanced Syntactic Test of Pronominal Reference (A-STOP). Limitations on testing time meant that it was only possible to give TAPS to 79 twin pairs (all 37 pairs from sample A and 42 pairs from sample B). A-STOP was given to the whole sample. The core language measures included the following.

The Test for Reception of Grammar (TROG) (Bishop, 1989) is a multiple-choice comprehension test in which the task is to select a picture to match a sentence spoken by the tester. All items use a simple vocabulary; grammatical complexity increases as the test proceeds. Four 4-choice items are used to test understanding of each of the 20 sentence types; a block of four items is scored as passed if all four items are responded to correctly and as failed if one or more errors occur.

Wechsler Comprehension, from the third revision of the Wechsler Intelligence Scale for Children (WISC-III) (Wechsler, 1992), measures a very different aspect of understanding – the ability to reply appropriately to questions in the absence of a concrete, pictorial context and to formulate a response to various hypothetical situations.

The Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals–Revised (CELF-R) (Semel, Wiig, & Secord, 1980) was used to assess expressive language. The testee repeats sentences of increasing grammatical complexity.

A word-finding test was devised to provide a measure of expressive vocabulary. A composite scale suitable for both adults and children was created from two existing tests: the Word-Finding Vocabulary Scale (Renfrew, 1980) and the Graded Naming Test (GNT) (McKenna & Warrington, 1983).

Three standardized tests were used to assess nonverbal reasoning ability: (1) Raven's Coloured Matrices or, for those older than 11½ years, Raven's Progressive Matrices (Raven, Court, & Raven, 1986), and (2) the Block Design and (3) Picture Completion subtests of the WISC-III. Scores on Block Design and Picture Completion were combined and prorated to give a short form estimate of performance IQ (PIQ).

Scores on Raven's Matrices and the four core language tests were converted to age-scaled scores, with a mean of 100 and a standard deviation of 15. Possible scores ranged from 55 to 145.

Test of Active and Passive Sentences

TAPS was developed by Van der Lely (1996) to assess comprehension of active and passive sentences. There were 12 core items, each consisting of four picture choices, as illustrated in the sample item in Figure 1. Two of the pictures illustrated an agent + action + theme/patient sentence, with different pictures illustrating the participants in both thematic roles (e.g., "man eats fish" and "fish eats man"). The remaining two picture choices depicted a theme/patient + state: that is, an adjectival passive interpretation (e.g., "the fish is eaten" and "the man is eaten"). In the course of the test, each set of four pictures was presented four times, once with each of the following sentences:

- (a) reversible active SVO (e.g., “the man eats the fish”);
- (b) reversible full passive (e.g., “the man is eaten by the fish”);
- (c) short progressive passive (e.g., “the fish is being eaten”); and
- (d) short passive with potentially adjectival passive interpretation (e.g., “the fish is eaten”).

For sentence types (a), (b), and (c), only one picture choice is correct. Sentence type (d) is ambiguous, and two interpretations are acceptable. The first is the verbal interpretation, corresponding to the picture where the agent carries out the action (e.g., picture 2 for “the fish is eaten”). The second is the adjectival passive interpretation (e.g., picture 3).

A constant random order was used for the 48 test items. For each item, the tester spoke the sentence and asked the child to find the picture that matched it. If a child hesitated between two choices or specifically commented on the ambiguity of sentences of type (d), the tester said, “Choose the one you think fits the sentence best.” Responses were categorized as follows:

- V: verbal passive interpretation with correct agent (e.g., for “the fish is eaten by the man,” picture 2 in Figure 1);
- R: verbal passive interpretation with agent/patient reversed (e.g., picture 4 in Figure 1);
- A: adjectival passive interpretation (only patient depicted) (e.g., picture 3 in Figure 1); and
- X: adjectival interpretation with wrong patient (e.g., picture 1 in Figure 1).

Advanced Syntactic Test of Pronominal Reference

We used a shortened version of the original test developed by Van der Lely and Stollwerck (1997). Each item consists of a picture paired with a spoken test sentence; the child is asked to respond “yes” or “no” depending on whether the picture matches the sentence. Pilot testing revealed that the full test was too long for most children, who became bored and restless, particularly so since this test was incorporated in a longer assessment battery. We therefore reduced the test from 128 to 48 items, omitting some sentence types where Van der Lely and Stollwerck had obtained ceiling effects (i.e., those where the two depicted characters were different genders) and reducing the number of items of each sentence type to 12. In the present article, we report data from the two sentence types that most directly relate to the RDDR hypothesis:

- NP: nominal main clause subject with personal pronoun (e.g., “Baloo Bear says Mowgli is tickling him”); and
- NR: nominal main clause subject with reflexive (e.g., “Baloo Bear says Mowgli is tickling himself”).

The 12 items for each sentence type included the following:

- Type M: for 4 items, the picture and test sentence matched (e.g., the picture in panel A of Figure 2 paired with “Baloo Bear says Mowgli is tickling him”).
- Type X: for 4 items, the picture and sentence did not match, and the correct agent performed the action but on the wrong patient (either self or other) (e.g., panel A of Figure 2 paired with “Baloo Bear says Mowgli is tickling himself” (NR-X) or panel B of Figure 2 paired with “Baloo Bear says Mowgli is tickling him” (NP-X)); and
- Type S: for 4 items, the picture and test sentence did not match, and the wrong agent carried out the stated action (e.g., panel A of Figure 2 paired with “Mowgli says Baloo Bear is tickling him” (NP-S) or panel B of Figure 2 paired with “Mowgli says Baloo Bear is tickling himself” (NR-S)).

A child who did not have any knowledge of Binding Principles might achieve a degree of success on items of type NR-X by recognizing that a reflexive refers to a self-directed action. This strategy, however, would not be effective for items of type NR-S.

RESULTS

Table 1 shows mean scores on the language and nonverbal composites for groups LI and LN and for children not meeting the criteria for LI or LN. As well as differing on the language variables that were used in defining categories, groups LI and LN differed in nonverbal ability and on age (for LI, $M = 10.75$, $SD = 1.69$; for LN, $M = 9.88$, $SD = 1.57$). Age and nonverbal ability were therefore included as covariates in group comparisons, where appropriate. We present results for TAPS, A-STOP, and TROG in turn and then consider the relationship between language deficits.

Test of Active and Passive Sentences

Overall accuracy. Data were available for 46 children in the LI group and 32 children in the LN group. A preliminary analysis was conducted to compare overall accuracy levels of groups LI and LN. Both adjectival passive (A) and verbal passive (V) responses were judged correct for the short ambiguous passives. There was a significant difference between groups: mean correct (out of 48) for group LI = 40.4 ($SD = 3.96$) and for group LN = 45.3 ($SD = 2.29$), $F(1, 76) = 39.8$, $p < .001$. Age was not significantly correlated with TAPS performance, $r(76) = -.047$, but nonverbal ability was: $r(76) = .420$ for Raven's Matrices and .445 for PIQ (both $p < .001$). The relationship between nonverbal ability and TAPS score did not differ significantly for the two groups. The analysis was therefore repeated with the average nonverbal score as a covariate. The group difference remained statistically significant, $F(1, 75) = 16.49$, $p < .001$. When performance is close to ceiling levels, as in the LN group, this is often seen as a problem, as it leads to reduced test sensitivity. However, such a

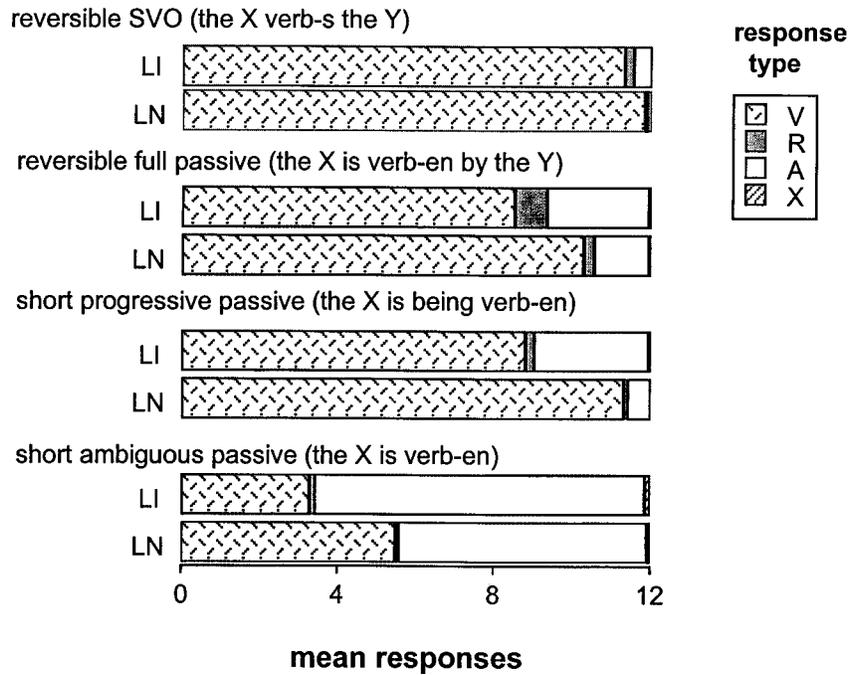


Figure 3. The mean proportion of responses of each kind for the four sentence structures in TAPS for children in groups LI and LN: V = verbal passive; R = reversed word order; A = adjectival passive; X = other.

pattern of results is consistent with Van der Lely’s description of G-SLI as a deficit in a grammatical module that is normally fully mature by around 7 years of age. The underlying model of disorder is categorical rather than continuous, and so we would expect control children of this age to have few difficulties on a test such as TAPS. For subsequent analyses, we defined syntactic deficit on TAPS as a total score of less than 42 correct. This corresponds roughly to the 5th centile for group LN; only one child in this group scored below this level. The majority (70%) of group LI did more poorly than this.

Response types. Figure 3 shows the mean proportion of responses of each kind for the four sentence structures for groups LI and LN. The pattern of performance is closely similar to that reported by Van der Lely (1996). For the short ambiguous passives, both adjectival and verbal interpretations were treated as correct. Performance on these sentences was close to ceiling for both LI and LN groups, which did not differ significantly; the mean for LI was 11.68 ($SD = 0.60$) and the mean for LN was 11.81 ($SD = .397$), $t(76) = 1.15$, $p = .236$. However, group LI showed a preference for the adjectival interpretation: the mean proportion of correct responses given an adjectival interpretation was .718 ($SD = .213$) for group LI and .536 ($SD = .217$) for group LN, $t(76) = 3.69$, $p < .001$. It

Table 2. Mean correct responses (out of four) for groups LI and LN on A-STOP

	LI (N = 54)		LN (N = 75)	
	M	SD	M	SD
<i>Name-Reflexive</i>				
Match	3.7	.57	3.8	.39
Mismatch-X	3.2	1.13	3.5	.83
Mismatch-S	3.0	.95	3.4	.86
<i>Name-Pronoun</i>				
Match	3.9	.29	3.9	.23
Mismatch-X	2.8	1.00	3.6	.71
Mismatch-S	2.2	1.31	3.1	1.10

is clear from Figure 3 that the most common error type for full and short progressive passive sentences was the adjectival interpretation for both groups LI and LN. Thus, for example, the picture of the eaten fish would be selected not only for “the fish is eaten” but also sometimes for “the fish is being eaten” or “the fish is eaten by the man”; for the latter two sentence types, this response was regarded as an error.

The other kind of response that is of interest in relation to the RDDR hypothesis is the reversal error, where the child selects a picture showing agent and patient but reverses the thematic role assignment. Errors of this kind were rare overall (see Figure 3) but were nevertheless significantly more common in group LI than in group LN (the median was zero for group LN and one for group LI). The statistical comparison used a nonparametric Mann-Whitney test because of nonnormality of the data, giving a value of $U = 533$, tied z value = 2.27, tied p value = .023.

Advanced Syntactic Test of Pronominal Reference

Overall accuracy. The mean number of items correct (out of four per condition) for the children in groups LI and LN is shown in Table 2. The mean total correct out of 24 was 18.72 ($SD = 2.90$) for group LI and 21.41 ($SD = 2.53$) for group LN; this difference was significant on a t test ($t = 5.61$, $p < .001$). The effect remained significant when age and nonverbal ability were included in ANCOVA as covariates, $F(1, 125) = 24.4$, $p < .001$. As with TAPS, the majority of LN children scored near ceiling (mean score of 89% correct, compared to 78% correct for group LI). A cutoff score of 17 corresponds to the 5th centile for group LN; in later analyses, this score was used to define syntactic deficit on this test. In group LI, 30% of children scored below this cutoff.

Error analysis. Van der Lely and Stollwerck (1997) argued that children with G-SLI tend to perform correctly on items where the sentence matches the picture because they generate an underspecified syntactic representation that is compati-

ble with the depicted state of affairs. The RDDR hypothesis predicts they will do especially poorly on NP-X, NP-S, and NR-S items, where they will accept a picture that does not match the sentence. Of course, errors on this test might also be made by a child who was just guessing at random, who was inattentive or failed to remember test sentences, or who had a response bias to say “yes.” We therefore computed the total number of errors on NP-X, NP-S, and NR-S items and then subtracted all other errors on A-STOP, to give an error index that reflected the extent to which the child’s errors were compatible with the RDDR hypothesis. This index gave a significant difference between groups LI and LN when the whole sample was considered, $t(127) = 4.96$, $p < .001$: for group LI, $N = 54$, $M = 2.83$, $SD = 1.89$; for group LN, $N = 75$, $M = 1.14$, $SD = 1.91$. It should be noted, however, that this “error pattern” score is not independent of accuracy because a child who makes no errors at all must obtain a difference score of zero. We performed a more stringent test to see whether a distinctive error pattern emerged when we restricted consideration only to those children who made at least one error. This meant excluding 16 children from the LN group, to give a mean of 1.46 ($SD = 2.05$). The LI versus LN comparison remained significant, $t(111) = 3.69$, $p < .001$.

Test for Reception of Grammar

Grammatical errors. TROG is usually scored in terms of the number of blocks that are passed, but on the earlier blocks of items it is possible to make errors that arise from selecting the wrong lexical item rather than because of grammatical problems. Other items test comprehension of meaning conveyed by prepositions and by inflections such as plural or comparative, which are not relevant for the RDDR hypothesis. TROG was therefore rescored in terms of the number of syntactic errors compatible with the RDDR hypothesis. For each child, we computed the total number of syntactic errors on blocks H (reversible active), L (reversible passive), N (subject postmodified by verb phrase or prepositional phrase), and R (object modified by relative clause). Block T (center-embedded sentences) was excluded because this block is usually failed by adults. The median number of errors for group LN was zero, and only two children in this group made more than three errors. The median number of errors for group LI was two (out of a maximum possible of 16). A Mann–Whitney test was used to compare the number of errors in the two groups and gave $U = 884$, $z = 5.52$, $p < .001$. Four or more errors was taken as the cutoff for defining syntactic impairment. Out of 54 children in the LI group, 18 (33%) had syntactic impairment on TROG.

Relationships among grammatical impairments

These analyses allow us to identify a set of possible indicators of G-SLI. We have three indicators based on overall scores on the grammatical comprehension tests – TAPS, A-STOP, and TROG. Each of these differentiated group LI from group LN. In addition, there are two qualitative indicators from TAPS – bias to

Table 3. *Frequency of G-SLI error types*

Classification	<i>N</i> ^a	Percentage showing error type					
		G1	G2	G3	G4	G5	G6
Language-impaired (LI)	54 (46)	70	63	33	30	57	33
Language normal (LN)	76 (32)	3	19	13	5	19	3
Marginal	12 (11)	64	73	9	17	67	8
Resolved	73 (39)	18	49	18	11	25	7
High IQ	35 (15)	13	60	0	0	9	0
Low IQ	32 (12)	42	33	33	34	44	25

Key: G1: errors on TAPS > 6 (i.e., accuracy below 5th centile for group LN); G2: bias to adjectival interpretation of short ambiguous passive (9 or more adjectival responses on 12 items; binomial probability of obtaining a score as extreme as this by chance is .073); G3: reversal errors on 2 or more TAPS items; G4: errors on A-STOP > 7 (i.e., accuracy below 5th centile for group LN); G5: A-STOP error difference score of 3 or more (total errors on NP-X, NP-S, and NR-S minus total errors on other item types); G6: grammatical errors on 4 or more items from TROG blocks H, L, N, and R.

^aThe number in parentheses denotes *N* tested on TAPS.

adjectival interpretation of ambiguous short passives and tendency to make reversal errors. These indicators also differentiated the groups. On A-STOP, a qualitative index reflecting the tendency to make disproportionate errors of the type predicted by the RDDR hypothesis (on items of type NP-X, NP-S, and NR-S) distinguished group LI from group LN. The next step is to consider whether these indicators are associated, which would enable us to identify a subset of children who fit the G-SLI subtype.

The entire sample of children, including those who did not meet the criteria for LI or LN, was used for these analyses. Table 3 shows the frequency of occurrence of each indicator in relation to language subgroup. Two important points emerge from this analysis. First, indicator G2 (bias to adjectival interpretation of ambiguous short passives) is of relatively high frequency in all groups except group LN. Particularly striking is the high frequency of this indicator in the small subgroup of children with nonverbal ability above 120. With hindsight, this is not surprising. A very bright child might detect the ambiguity of these sentences and develop a strategy for responding to them in a particular way. Indeed, some children did specifically point out that two responses were possible and asked which one they should select. It would seem, then, that an adjectival bias on ambiguous short passives can arise for different reasons; although in some children it might indicate a failure to compute the complex passive interpretation, in others it might be evidence of a sophisticated strategic approach to the task. The second point to note is the relatively high level of all indicators in children of low nonverbal ability. Many of these children did have clear evidence of language difficulties. However, if we include such children in our analyses, it becomes difficult to separate specific grammatical difficulties

Table 4. *Correlations (above diagonal) and phi coefficients (below diagonal) for associations between quantitative and categorical indicators of G-SLI (excluding low IQ group)*

	G1	G2	G3	G4	G5	G6
G1	•	.309***	.473***	.335***	.303***	.414***
G2	.160	•	(-.158)	.152	(-.133)	(-.096)
G3	.398***	(.043)	•	.445***	.233**	.352***
G4	.222*	(.174)	.406***	•	.571***	.495***
G5	.267**	(-.093)	.208*	.363***	•	.360***
G6	.277**	(.028)	.234**	.151*	.149*	•

Note: $N = 144$ for associations with variables G1 to G3 and $N = 254$ for associations with the other variables.

* $p < .05$; ** $p < .01$; *** $p < .001$ (only positive associations are marked; negative associations, i.e., opposite to prediction, shown in parentheses).

from more general problems with attention and memory. Therefore, when studying relationships between indicators, these children were excluded so we could focus attention on those of average nonverbal ability.

Table 4 shows relations between indicators. Phi coefficients, which measure categorical association, are shown below the diagonal. These reflect how far a child's impairment status on one indicator is associated with impairment on another. Correlations are shown above the diagonal: these are based on the whole range of raw test scores from which the indicators are derived, without using any cutoff. In general, the strongest associations are among different indices from the same test (i.e., G1–G3 and G4–G5). Logically, G1 and G3 and G4 and G5 are not fully independent, as the child who makes very few errors on TAPS or A-STOP would be bound not to fall beyond cutoff on G3 and G5, respectively. The three overall deficit indicators from TAPS (G1), A-STOP (G4), and TROG (G6) are significantly related on both quantitative and categorical measures of association, though correlations are well below 1.0, indicating that the association is far from perfect. The index of adjectival bias (G2) is not significantly associated with other indicators apart from G1. This presumably reflects the fact that this bias can arise for different reasons, as suggested earlier. For this reason, G2 was dropped from subsequent analyses, leaving five indicators of G-SLI.

After excluding the low IQ group, full data on all indices were available for 144 children, 2 children scored positive on all five indicators, and 9 children scored positive on four of the five indicators. These numbers are far too small for genetic analysis, but Table 5 shows the characteristics of affected twins and their cotwins. If we take a score of 4 or more indicators of G-SLI to define affected status, only a fifth of the pairs of MZ twins and a quarter of the pairs of DZ twins are concordant. Cotwins of affected children did tend to score positive on some of the indicators of G-SLI without meeting our criteria for G-SLI.

It should be noted that the children who met our criterion for G-SLI were

very different in their nongrammatical characteristics from AZ, the paradigmatic case described by Van der Lely (1997). A noteworthy finding was that three of the six G-SLI cases from sample A demonstrated an appreciable drop in nonverbal ability from their initial assessment to the reassessment for this study. Had we relied solely on the most recent measure of nonverbal ability to define LI status, these children would have been excluded because of low IQ. Furthermore, in the majority of cases of G-SLI, deficits were seen in the nongrammatical language measures such as the word-finding test. At the other extreme, among the DZ twins, some of the cases that met the criteria for G-SLI were not regarded as language-impaired on the basis of either standardized tests or therapy status.

DISCUSSION

This study replicated previous research by Van der Lely and colleagues, showing that two comprehension tests that required the child to compute syntactic dependencies discriminated between children with language impairments and those with normal language. On the test of active and passive sentences, we found that the pattern of responses was closely similar to that previously reported by Van der Lely (1996). However, although there were significant associations between these different indicators of deficit, these were far from perfect, and very few children in the LI group showed a pattern of impairment consistent with the RDDR hypothesis on all indicators. When we identified those few children who did score positive on four or five indicators of G-SLI, they did not resemble the cases previously described by Van der Lely and colleagues in terms of the specificity of their impairment. It was unfortunate that there were too few children who might be regarded as cases of G-SLI to allow for meaningful genetic analysis.

Cases such as AZ (Van der Lely, 1997), where syntactic skills are more seriously affected than other aspects of language and there is good nonverbal ability, are of considerable theoretical importance in demonstrating the independence of grammar from other cognitive functions. Van der Lely and colleagues (1998) described six children with SLI who showed a similar pattern of selective impairment in grammatical language tests, with much better performance on nongrammatical language tests, and who performed well on nonverbal measures. Clearly, AZ is not unique. However, our data suggest that this highly selective profile of language impairment is relatively rare.

Children in this sample differed from AZ in two important respects. First, those who had the comprehension difficulties predicted by the RDDR hypothesis tended also to have major impairments on nonsyntactic language tasks (e.g., word-finding) and/or limited nonverbal ability. It is not entirely surprising to find that children with syntactic difficulties tend to have limited vocabulary. As Bishop (1997b) argued, a child who has a restricted understanding of grammar will be hampered in using syntactic bootstrapping to deduce the meanings of new words (see also O'Hara & Johnston, 1997; Van der Lely, 1994). Nevertheless, AZ and the other cases of G-SLI studied by Van der Lely and colleagues managed to attain relatively good levels of vocabulary despite their grammatical

Table 5. Characteristics of twins where one or both score positive on four or more indicators of G-SLI

ID ^a	MZ/DZ	Age	Sex	G-SLI	TROG	RS	COMP	WF	RM	PIQ	Phonology impaired?	Notes ^b
A104A	MZ	11.42	F	4	66	62	80	64	75	62	Never	Therapy for 4 years from age 2.5 Raven's Matrices IQ = 87 at age 7
A104B	MZ	11.42	F	3	69	61	88	92	75	81	Never	Therapy for 4 years from age 2.5 Raven's Matrices IQ = 109 at age 7
A144A	MZ	10.08	M	5	69	55	65	62	80	64	Yes, severe	Therapy from age 2.5; now in language unit Raven's Matrices IQ = 85 at age 7
A144B	MZ	10.08	M	3	66	57	72	55	78	59	Yes, severe	Therapy from age 2.5; now in language unit Raven's Matrices IQ = 81 at age 7
A154A	MZ	10.58	M	4	75	69	85	82	92	106	Resolved	Therapy for 2 years from age 3; then remedial help at school
A154B	MZ	10.58	M	4	83	65	80	67	77	70	Resolved	Therapy for 2 years from age 3; then remedial help at school Raven's Matrices IQ = 103 at age 7
A161A	MZ	10.00	M	4	94	85	110	93	96	95	Never	Never any therapy Dysfluent at age 10
A161B	MZ	10.00	M	2	103	99	65	88	90	72	Resolved	Therapy for 6 months from age 5 Otitis media plus surgery Dysfluent at age 10
B007A	MZ	8.46	F	4	83	75	70	81	86	87	Resolved	Preschool therapy for about 1 year, plus some extra help in school
B007B	MZ	8.46	F	3	103	87	80	83	93	86	Resolved	Preschool therapy for about 1 year, plus some extra help in school Hospital referral for persistent otitis media

A145A	DZ	12.17	M	0	111	89	112	95	89	98	Never	No parental concern or therapy
A145B	DZ	12.17	M	4	67	77	85	66	70	79	Highly unintelligible; diagnosed as dyspraxic	Therapy for 4 years from age 3; special education placement Raven's Matrices IQ = 87 at age 8 Hospital referral for otitis media but no surgery
B027A	DZ	7.67	F	3	86	101	70	99	80	84	Never	Never any therapy or parental concern Resuscitated at birth; special care at 5 weeks Otitis media plus surgery
B027B	DZ	7.67	F	4	99	94	85	107	90	102	Never	Never any therapy or parental concern Resuscitated at birth; special care at 5 weeks
B028A	DZ	9.58	M	2	87	87	100	67	96	91	Never	Never any therapy or parental concern In special care baby unit for 2 weeks Mild otitis media but no hospital referral
B028B	DZ	9.58	M	5	82	79	110	77	88	106	Never	Assessed for learning difficulties but never any in- tervention In special care baby unit for 2 weeks Mild otitis media but no hospital referral
B303A	DZ	7.50	F	4	94	94	65	96	114	117	Never	History of brief therapy only as a "marginal case" Mild otitis media but no hospital referral
B303B	DZ	7.50	F	4	90	81	60	100	93	95	Never	As for cotwin

Key: G-SLI: number of positive markers for G-SLI; RS: Recalling Sentences; COMP: Wechsler Comprehension; WF: Word-Finding; RM: Raven's Matrices IQ; PIQ: short form Wechsler PIQ.

^aInitial letter denotes whether sample A or B. Children with same number are members of same twin pair. Those meeting criteria for G-SLI are shown in boldface.

^bIncludes information on IQ decline in sample A and on perinatal, medical, and intervention history in both samples.

limitations. As far as nonverbal ability is concerned, in our sample of children with G-SLI there were several with nonverbal ability well within the normal range; however, none came close to the high levels of ability seen in AZ. In addition, there were some children who had achieved nonverbal IQs in the normal range when previously assessed but who scored in the impaired range when retested for this study. Again, this is perhaps not surprising: as Leonard (1998) noted in a review of the field, many children with SLI show a decline in nonverbal ability with age. He suggested that many cognitive skills depend on verbal mediation even if they do not require overt language, which means that children with language limitations will suffer.¹ However, the cases of G-SLI described by Van der Lely et al. (1998) demonstrate that it is possible to have marked discrepancies between poor grammar and good nonverbal abilities, and it is intriguing to consider why some children manage to do well on reasoning tasks despite their lack of verbal ability.

The fact that our cases with G-SLI had additional deficits outside the syntactic system could indicate that the underlying neurobiological impairment was more extensive in these children. We considered whether other factors in the child's medical history might account for the associated deficits seen in our sample. One issue to consider is whether twins might have more widespread difficulties because they experience a relatively high rate of perinatal problems. This seems unlikely. We had gathered data on perinatal status and medical history on this sample, both from parental report and medical records, and found no evidence of any link with language impairment (Bishop, 1997a). Any indications of abnormal perinatal or medical history are reported in the final column of Table 5. Most children had uneventful histories. As in the sample studied by Van der Lely and Stollwerck (1996), where children with major perinatal problems or neurological risk factors had been excluded, a family history of language impairment emerged as the main risk factor for language impairment. This does not, of course, mean that the same etiological processes are involved in all cases: there may be a range of genetic defects that can lead to a final common path of grammatical impairment, some of which cause widespread deficits and others which have more selective effects.

Another factor that might influence children's profiles is remediation. All of the cases of G-SLI studied by Van der Lely were recruited from special residential schools for children with SLI. As can be seen from Table 5, only two cases of G-SLI in the current sample were receiving full-time special education. It may be that intensive speech and language therapy builds the child's nonsyntactic language skills so that a core underlying impairment in grammar is revealed more starkly.

Although our failure to find pure cases in this sample does not challenge the concept of a specific form of grammatical language impairment, it does raise questions about how far we can generalize from cases of G-SLI to the SLI population as a whole. It is worth noting that, although we found very few children with all the grammatical characteristics of G-SLI, we did find a substantial number of children who did poorly on some of the indicators of this disorder. Our finding of children with what appears to be partial indications of G-SLI poses problems for the RDDR, which treats G-SLI as an all-or-none

condition in which a modular grammatical system has failed to develop properly. Either we must assume these children do have G-SLI but for some reason do not show the full range of features of this disorder, or we must conclude that G-SLI only affects a small minority of language-impaired children, but that similar grammatical problems can arise for quite different reasons in other children.

Let us consider the first possibility, that G-SLI might manifest in a partial form. This could occur if the underlying modular impairment in G-SLI is not static. Rather than being permanently impaired, grammatical knowledge could be simply very slow in developing so that, as they grow older, at least some children with G-SLI develop the ability to process more distant syntactic dependencies. Although compatible with some maturational accounts of syntax (e.g., Wexler, 1990), this view appears to be at odds with the data from the G-SLI cases studied by Van der Lely. Many of these individuals are now in their late teens and early twenties. However, they continue to have grammatical problems and make grammatical errors similar to those they made in childhood. An alternative view is that a child with G-SLI might achieve success on a task such as TAPS by learning cognitive strategies to interpret and produce sentences, especially if there has been a long history of speech–language therapy. The appearance of grammatical abilities many years after they normally develop may not reflect the use of the same mechanisms or representations as are seen in normally developing children. However, if this is so, it should be possible to devise test materials so that different reasons for success or failure can be distinguished in terms of response latency or error profile. A final possibility is that deficits within the grammatical system may fractionate. There is, for instance, evidence that morphological representations of regular words are impaired in some but not all children with SLI (Oetting & Horohov, 1997; Oetting & Rice, 1993; Van der Lely & Christian, 2000).

The second way of interpreting our data is to suppose that some children with SLI make errors on grammatical comprehension tests for reasons other than difficulties in computing syntactic relationships. Perceptual limitations and restricted processing capacity do not seem plausible explanations for the deficits in cases such as AZ (Van der Lely et al., 1998), but these factors could be implicated in other children (Leonard, 1998). For instance, on a test such as TAPS the child might fail to remember the test sentences, have lapses of attention and so fail to take in critical information, or misperceive critical information. For instance, the child might mishear “the fish is being eaten by the man” as “the fish has been eaten by the man” and so select the adjectival rather than the verbal interpretation, which would constitute an error on this test. On A-STOP, a lack of sensitivity to prosody could lead the child to treat the sentences as ambiguous (“Baloo Bear, says Mowgli, is tickling him” vs. “Baloo Bear says Mowgli is tickling him”).

In conclusion, tests of understanding of syntactic dependencies are sensitive to the language difficulties of children with SLI. However, the highly selective and consistent pattern of deficits described in G-SLI is relatively rare. It remains unclear whether the SLI population contains a mixture of children who fail such tests for different reasons or whether there is a fractionation of components of

the grammatical system in SLI. It is only through a careful analysis of grammatical abilities across a broad range of children that we can start to address this issue.

ACKNOWLEDGMENTS

Thanks are due to Tom Delaney for assistance with testing, to Sandra Stops for scheduling and data coding, and to Dr Tony Monaco for assistance with zygoty testing.

This study would not have been possible without the generous cooperation of the education authorities and schools in Hertfordshire and Cambridgeshire. We also extend our warmest thanks to the families of twins who participated.

NOTE

1. We need also to consider the possibility that the decline could reflect measurement error. If an early intelligence test gave spuriously good results, then the child might appear to have a big discrepancy between verbal and nonverbal skills. However, on retest, we would expect to see regression to the true score, which would look like a decline in IQ.

REFERENCES

- Adams, C. (1990). Syntactic comprehension in children with expressive language impairment. *British Journal of Disorders of Communication*, 25, 149–171.
- Bishop, D. V. M. (1979). Comprehension in developmental language disorders. *Developmental Medicine and Child Neurology*, 21, 225–238.
- (1982). Comprehension of spoken, written, and signed sentences in childhood language disorders. *Journal of Child Psychology and Psychiatry*, 23, 1–20.
- (1989). *Test for Reception of Grammar* (2nd ed.). Age and Cognitive Performance Research Centre, University of Manchester.
- (1997a). Pre- and perinatal hazards and family background in children with specific language impairments: A study of twins. *Brain and Language*, 56, 1–26.
- (1997b). *Uncommon understanding: Development and disorders of language comprehension in children*. Hove: Psychology Press.
- Bishop, D. V. M., Bishop, S. J., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language and Hearing Research*, 42, 155–168.
- Bishop, D. V. M., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: Evidence from a twin study. *Developmental Medicine and Child Neurology*, 37, 56–71.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.
- Lewis, B. A., & Thompson, L. A. (1992). A study of developmental speech and language disorders in twins. *Journal of Speech and Hearing Research*, 35, 1086–1094.
- McKenna, P., & Warrington, E. K. (1983). *Graded Naming Test (GNT)*. Windsor: NFER-Nelson.
- Nichols, R. C., & Bilbro, W. C. (1966). The diagnosis of twin zygoty. *Acta Geneticae Medicae et Gemellologiae*, 16, 265–275.
- Oetting, J. B., & Horohov, J. E. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language and Hearing Research*, 40, 62–74.
- Oetting, J. B., & Rice, M. L. (1993). Plural acquisition in children with specific language impairment. *Journal of Speech and Hearing Research*, 36, 1236–1248.
- O'Hara, M., & Johnston, J. (1997). Syntactic bootstrapping in children with specific language impairment. *European Journal of Disorders of Communication*, 32, 189–205.
- Raven, J. C., Court, J. H., & Raven, J. (1986). *Raven's Progressive Matrices and Raven's Coloured Matrices*. London: H. K. Lewis.

- Renfrew, C. (1980). *Word-Finding Vocabulary Scale*. Oxford: Author.
- Rosen, S., Van der Lely, H. K. J., & Dry, S. (1997). Speech and nonspeech auditory abilities in two children with disordered language. *UCL Department of Phonetics: Work in Progress*, 10, 186–198.
- Semel, E. M., Wiig, E. H., & Secord, W. (1980). *Clinical Evaluation of Language Fundamentals – Revised*. San Antonio, TX: Psychological Corporation.
- Tomblin, J. B., & Buckwalter, P. R. (1994). Studies of genetics of specific language impairment. In R. Watkins & M. Rice (Eds.), *Specific language impairments in children* (pp. 17–34). Baltimore: Paul H. Brookes.
- Van der Lely, H. K. J. (1990). *Sentence comprehension processes in specifically language impaired children*. Unpublished doctoral thesis, Birkbeck College, University of London.
- (1994). Canonical linking rules forward versus reverse linking in normally developing and specifically language-impaired children. *Cognition*, 51, 29–72.
- (1996). Specifically language impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua*, 98, 243–272.
- (1997). Language and cognitive development in a grammatical SLI boy: Modularity and innateness. *Journal of Neurolinguistics*, 10, 75–107.
- (in press). SLI in children: Movement, economy and deficits in the computational-syntactic system. *Language Acquisition*.
- Van der Lely, H. K. J., & Christian, V. (2000). Noun compound formation in specifically language impaired children. *Cognition*, 75, 33–63.
- Van der Lely, H. K. J., & Harris, M. (1990). Comprehension of reversible sentences in specifically language impaired children. *Journal of Speech and Hearing Research*, 55, 101–117.
- Van der Lely, H. K. J., Rosen, S., & McClelland, A. (1998). Evidence for a grammar-specific deficit in children. *Current Biology*, 8, 1253–1258.
- Van der Lely, H. K. J., & Stollwerck, L. (1996). A grammatical specific language impairment in children: An autosomal dominant inheritance. *Brain and Language*, 52, 484–504.
- (1997). Binding theory and grammatical specific language impairment in children. *Cognition*, 62, 245–290.
- Wechsler, D. (1992). *Wechsler Intelligence Scale for Children – Third UK edition*. London: Psychological Corporation.
- Wexler, K. (1990). Innateness and maturation in linguistic development. *Developmental Psychobiology*, 23, 645–660.