

Riches NG.

[Complex sentence profiles in children with Specific Language Impairment: are they really atypical?](#)

Journal of Child Language (2016)

DOI: <http://dx.doi.org/10.1017/S0305000915000847>

Copyright:

This is the author's accepted manuscript of an article that has been published in its final definitive form by Cambridge University Press, 2016.

DOI link to article:

<http://dx.doi.org/10.1017/S0305000915000847>

Date deposited:

14/09/2015



This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#)



Complex sentence profiles in children with Specific Language Impairment: are they really atypical?

Journal:	<i>Journal of Child Language</i>
Manuscript ID:	JCL-08-13-073.R4
Manuscript Type:	General Article
Keywords:	syntax, specific language impairment

SCHOLARONE™
Manuscripts

For Review

Repetition of complex sentences in SLI

ABSTRACT

Children with Specific Language Impairment (SLI) have language difficulties of unknown origin. Syntactic profiles are atypical, with poor performance on non-canonical structures, e.g. object relatives, suggesting a localised deficit. However, existing analyses using ANOVAs are problematic because they do not systematically address unequal variance, or fully model random effects. Consequently, a Generalised Linear Model (GLM) was used to analyse data from a Sentence Repetition (SR) task involving relative clauses. 17 children with SLI (mean age 6;7), 21 Language Matched (LM) children (mean age 4;8) and 17 Age Matched (AM) children (mean age 6;5) repeated 100 canonical and non-canonical sentences. ANOVAs found a significant Group by Canonicity interaction for the SLI versus AM contrast only. However, the GLM found no significant interaction. Consequently, arguments for a localised deficit may depend on statistical methods which are prone to exaggerate profile differences. Nonetheless, a subgroup of SLI exhibited particularly severe structural language difficulties.

Repetition of complex sentences in SLI

INTRODUCTION

About 7% of children experience unexplained language difficulties, a condition commonly referred to as Specific Language Impairment (SLI) (Tomblin, Records, Buckwalter, & Zhang, 1997). Language difficulties are not due to factors normally associated with poor language; low IQ, neurological damage, hearing difficulties, or other known syndromes such as autism. While many language subdomains are impaired, it is often proposed that morphosyntax is most severely affected (Leonard, 2000). For example, the expressive language of children with SLI is characterised by grammatical errors involving the omission of tense /agreement / aspect morphemes, and incorrect case-marking, e.g. *He drops it* → *him drop it*, *She is sleeping now* → *her sleep now*. Nonetheless, difficulties are also evident in other language subdomains such as vocabulary, phonology, and pragmatics.

A central question in SLI research is whether these children are delayed, or whether their language is ‘deviant,’ or qualitatively different to that of typically-developing children. To adopt the terminology of Leonard (2014: 94), qualitative differences may be observed at the ‘macro’ and ‘micro’ levels. The macro level refers to performance across different language subdomains, e.g. morphosyntax, phonology, and vocabulary, while the micro level refers to performance within a subdomain, e.g. verb versus noun morphology. Studies often refer to differences in PROFILES, which is a pattern of performance across different language assessments. At the macro level, children with SLI perform worse on tests of morphosyntactic abilities than tests of lexical knowledge, compared with typically developing children who exhibit similar performance across these two subdomains (Rice, Wexler, & Cleave, 1995). At the micro-level, within the subdomain of morphology, verb affixes, e.g. *watch-ed*, are more prone to omission than noun affixes, e.g. *dog-s*, whereas profiles are flatter in typically-developing children (e.g. Rice,

Repetition of complex sentences in SLI

Wexler, & Cleave, 1995). Complex sentences also reveal idiosyncratic profiles. For example, children with SLI perform much better on subject relatives (1) than object relatives (2) (e.g. Novogrodsky & Friedmann, 2006);

(1) The dog that chased the cat was brown SUBJECT RELATIVE

(2) *The cat_i that the dog chased t_i was ginger* OBJECT RELATIVE

Again, age- and language-matched groups do not exhibit such a differential. Unusual profiles are apparent at even lower levels of granularity. For example, with regard to past tense production, a subdomain of morphology (or a sub-sub-domain), children with SLI have a specific difficulty with regular past tense, but are comparatively good on irregulars, a profile less pronounced in typically-developing children (van der Lely & Ullman, 2001).

Longitudinal studies complement the findings of cross-sectional data. While in typically-developing children, language skills develop in an integrated fashion, growth curves in SLI are characterised by ‘islands’ of extreme delay. Whereas vocabulary is often delayed, morphosyntax lags yet further behind vocabulary, and within morphosyntax, tense-marking appears most severely delayed (Rice, 2013). Consequently, tense-marking difficulties have been characterised as a ‘delay-within-a-delay’.

While uneven language profiles do not play a role in the diagnosis of SLI, for many researchers they constitute a defining characteristic. For example, Leonard (2014: 94) argues that “a profile difference appears to be the most accurate [way of characterising SLI] both at a macro and micro level”. Without such profile differences, it would be difficult to motivate SLI as a distinct diagnostic category, as opposed to a term to describe children at the tail end of the normal distribution. Excluding the possibility that these children lie at the low end of the language

Repetition of complex sentences in SLI

1
2
3 continuum is an essential first step to developing causal theories of SLI, e.g. theories proposing
4
5 specific grammatical deficits, as proposed by De Villiers (2003). In addition, profile differences
6
7 have motivated numerous causal theories of SLI, including the Extended Optional Infinitive
8
9 account (Rice, Wexler, & Cleave, 1995) which addresses morphosyntactic profiles, and accounts
10
11 of difficulties with complex sentences (see below). However, there is a striking disconnect
12
13 between experimental investigations of language profiles, which suggest that children with SLI
14
15 have qualitatively different language systems, and large-scale epidemiological / taxonomic
16
17 studies which find little support for the claim that these children should be regarded as belonging
18
19 to a distinct category (Dollaghan, 2004, 2011; Tomblin & Zhang, 2006). Partly in response to this,
20
21 many researchers are now openly questioning SLI as a diagnosis (Reilly et al., 2014)
22
23
24
25
26
27

28 This paper investigates profiles during complex sentence production, and therefore it
29
30 focuses on the micro level. It also focuses on cross-sectional as opposed to longitudinal data.
31
32 While there is no single agreed-upon definition of linguistic complexity, it is often argued that
33
34 sentences are more complex when they involve NON-CANONICAL word order. Object relatives (2)
35
36 are non-canonical in placing the Object before the Subject. Object questions (4) also exhibit the
37
38 same property;
39
40
41
42

43 (3) Which dog was chasing the cat? SUBJECT QUESTION
44

45
46 (4) *Which dog_i was the cat chasing t_i ?* OBJECT QUESTION
47
48

49 While CANONICITY is not the only way to define complexity, it is relatively easy to manipulate, and
50
51 consistently affects processing difficulty. Typically developing children find non-canonical
52
53 structures more difficult to understand / produce than canonical structures (e.g. Tyack & Ingram,
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 1977), and numerous studies have demonstrated this discrepancy is even greater in children with
4
5 SLI (see below).
6
7

8
9 Accounts of difficulties with complex sentences may be roughly divided into those
10
11 proposing an underlying difficulty with linguistic competence, and those which suggest deficient
12
13 processing mechanisms. While early linguistic accounts proposed difficulties with long distance
14
15 relationships (van der Lely & Battel, 2003), more recent accounts have suggested a difficulty with
16
17 thematic role assignment in the context of sentences with long-distance movement (Friedmann &
18
19 Novogrodsky, 2007). In particular, children with SLI may operate with a strict version of Rizzi's
20
21 (1990) Relativised Minimality (Friedmann & Novogrodsky, 2011). This outlaws complex
22
23 sentences where a constituent crosses over a constituent of the same type, e.g. referential NP (*the*
24
25 *cat*) and referential NP (*the dog*) in (2) and (4).
26
27
28
29
30

31 Processing accounts address the same structures from the perspective of capacity
32
33 limitations. In order to interpret (2) and (4) children must store the displaced NP until it can be
34
35 thematically integrated at the trace. This process involves maintaining a phonological / lexical
36
37 representation whilst processing the remainder of the sentence. In this way, the sentence places a
38
39 greater burden on verbal working memory (WM) than the corresponding canonical alternative.
40
41 Many researchers have argued that storage and processing compete for limited resources (Just &
42
43 Carpenter, 1992), and consequently maintaining displaced NPs will detract from processing in
44
45 non-canonical sentences. In addition, NPs with similar characteristics may interfere with each
46
47 other in verbal WM, and such interference effects are greater in non-canonical sentences where
48
49 displaced NPs must be must be maintained while processing the intervening NP (Gordon,
50
51 Hendrick, & Johnson, 2001; see also Gibson & Pearlmutter, 1998, for a concise introduction). In
52
53 addition to the processing costs involved in movement, relative clauses also involve perspective
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

switching, whereby the head of the relative clause has two thematic roles; one with respect to the relative clause, and another with respect to the main clause. For example, in (2) *cat*, is the patient / object of *chase*, but the subject of *be*. Again, perspective-switching may tax WM resources (Booth, MacWhinney, & Harasaki, 2000). In summary, complex sentence interpretation depends on WM abilities, which are often compromised in SLI. Consequently, there may be a causal relation between WM limitations in SLI and their difficulties understanding / producing complex sentences (Montgomery, Magimairaj, & Finney, 2010).

The analysis of profiles

Profile differences are of major theoretical importance in SLI research, contributing to the claim that language in SLI is disordered as opposed to weak. Typically, in cross-sectional data, profiles are explored by investigating the interaction between Group (between-subjects factor) and linguistic construct (within-subjects factor), where the linguistic construct is the type of morpheme (e.g. noun affix versus verb affix), or canonicity. If the interaction term is significant researchers conclude that profiles are different across groups. Interactions are always ‘quantitative’ with the effect of condition having the same polarity across groups. For example, all groups find non-canonical sentences more difficult than canonical sentences, but the magnitude of this effect varies between groups.

Interactions between Group and Canonicity have been investigated by most studies of SLI assessing comprehension and production of relative clauses and questions across a variety of languages including English, Hebrew, Greek, Italian, Danish and Cantonese. The majority have identified a significant Group by Canonicity interaction on at least one dependent measure (Adani, Forgiarini, Guasti, & van der Lely, 2013; Deevy & Leonard, 2004; Friedmann & Novogrodsky, 2007; Jensen de López, Olsen, & Chondrogianni, 2012; Novogrodsky & Friedmann, 2006; Riches,

Repetition of complex sentences in SLI

1
2
3 Loucas, Charman, Simonoff, & Baird, 2010; van der Lely & Battel, 2003; van der Lely, Jones, &
4
5 Marshall, 2011). A Group by Structure interaction was also observed by Wong, Leonard, Fletcher
6
7 and Stokes (2004) who investigated question formation in Cantonese children with SLI, and found
8
9 that they had particular difficulties with object questions. However, in Cantonese these are in fact
10
11 canonical structures, exhibiting the same word order as declarative sentences with a transitive
12
13 verb, and consequently the authors provide a non-movement account based on input frequency and
14
15 animacy constraints. The above list does not do justice to the complexity of these studies and their
16
17 detailed investigation of different error types, e.g. mis-assignment of thematic roles (Friedmann &
18
19 Novogrodsky, 2007). It also overlooks finegrained distinctions in the structures used, e.g.
20
21 right-branching relatives in Novogrodsky & Friedmann (2006) and cross-linguistic differences.
22
23 However, it is accurate to say that when errors / items correct are counted, they all identify a
24
25 distinct profile in children with SLI with regard to non-canonical structures. By contrast, a few
26
27 studies do not identify significant Group by Canonicity interactions (Epstein, Hestvik, Shafer, &
28
29 Schwartz, 2013; Stavrakaki, 2001). However, despite their findings, the researchers do not
30
31 question the assumption that profile differences exist, probably due to strong converging evidence.
32
33
34
35
36
37
38

39
40 With the exception of Adani et al. (2013) all of the above studies use two-way ANOVAs
41
42 to analyse count data (number of items correct / incorrect). Sometimes counts are expressed as
43
44 percentages, but this makes no difference to the statistical results, as the ratios between
45
46 observations are unaltered. Often, typically-developing children perform close to ceiling (e.g. 94%
47
48 accuracy on object relatives in Novogrodsky and Friedmann, 2006, and 87% for object questions
49
50 in Deevy and Leonard, 2004). This can be regarded as an occupational hazard when comparing
51
52 profiles where groups differ greatly in ability. If a group performs close to ceiling the VARIANCE of
53
54 the data (roughly-speaking the spread of data points around the mean) is reduced. This
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 phenomenon is sometimes referred to as a ‘ceiling effect’, and is generally regarded as problematic
4
5 for statistical analysis. This is because if variance differs greatly across different groups and / or
6
7 conditions, then the RESIDUALS (roughly-speaking the difference between observed and predicted
8
9 values) will depart from a normal distribution. This will affect significance tests, e.g. the t-test,
10
11 z-test and F-test, as they all assume normally distributed residuals.
12
13
14
15

16 One way to minimise the impact of a ceiling effect is to use a transformation, e.g. arc
17
18 sine, square root or log transformation. These ensure that variances do not vary greatly across
19
20 groups and / or conditions, i.e. they are ‘stabilised’. An alternative is to use an F-test which is
21
22 robust to heterogeneity of variance, e.g. Games-Howell (Field, 2000, p. 276). However, there are
23
24 no reliable and universally-accepted guidelines to determine when to transform and how to
25
26 transform, or which F-test to use. Moreover, studies vary in their treatment of the data, with some
27
28 studies applying transformations to count data (e.g. Deevy & Leonard, 2004; Wong et al., 2004)
29
30 and others (e.g. Friedmann & Novogrodsky, 2011; Stavrakaki, 2001) leaving count data
31
32 untransformed.
33
34
35
36
37

38 An alternative to the ANOVA is to use a GENERALISED LINEAR MODEL (GLM), a form
39
40 of regression analysis. This is distinct from the ‘General’ Linear Model which underlies the
41
42 ANOVA. GLMs extend or ‘generalise’ the basic linear model so that it deals with different types
43
44 of distribution. They do this by employing different ‘link functions’ (e.g. the log function for
45
46 modelling count data), and assuming different distributions (e.g. the poisson distribution which
47
48 tends to arise from count data; see Howell, 2013 for details). Mathematically-speaking there is no
49
50 difference between ANOVAs and simple linear models. However, GLMs allow greater flexibility
51
52 for dealing with data from different distributions, and crucially they allow for much better
53
54 modelling of residuals resulting from unequal variances. They do this systematically, as different
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 types of GLMs are specifically-designed for different distributions. By comparison, when using
4 ANOVAs, procedures for dealing with unequal variances are difficult to apply in a systematic
5 fashion. In addition, there is evidence to suggest that GLMs give more reliable results. For
6 example, Jaeger (2008) and Dixon (2008), compared the findings of ANOVAs on arc sine
7 transformed data with logistic regressions designed to model dichotomous data. Jaeger
8 investigated data on relative clause comprehension, later published by Arnon (2010), and Dixon
9 conducted a simulation study. Both studies found that ANOVAs performed poorly when data
10 approached the extremes in comparison to the GLM. Jaeger identified a bias towards a significant
11 interaction, while Dixon argued that the bias could increase or decrease the chances of a significant
12 interaction depending on the shape of the data. Together these studies suggest that GLMs may be a
13 better means of analysing interaction effects than ANOVAs where unequal variances arise.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

30 A further advantage of GLMs is that one can control for both by-items and by-subjects
31 effects using mixed effects models (Baayen, Davidson, & Bates, 2008). Modelling by-items
32 effects is important in ensuring that findings generalise beyond the current set of items (Clark,
33 1973). Furthermore, a failure to fully model random effects can lead to a Type I error, the incorrect
34 rejection of the null hypothesis (Barr, Levy, Scheepers, & Tily, 2013; Clark, 1973; Quené & Van
35 den Bergh, 2008). However, ANOVAs are unable to do this as they cannot simultaneously model
36 by-items and by-participant effects. For example, by-subjects ANOVAs ‘aggregate’ data at the
37 participant level by ensuring that there is only one observation per participant for each cell of the
38 design.
39
40
41
42
43
44
45
46
47
48
49
50
51

52 These are abstract statistical issues which are rarely discussed in child language
53 research. However, they relate directly to theory and the way we conceptualise SLI. Numerous
54 researchers have argued for localised syntactic difficulties based on group by linguistic construct
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 interactions. However, ANOVAs may not be the best procedure for testing these claims. Only one
4
5 previous study by Adani and colleagues (2013) has employed a GLM to analyse complex sentence
6
7 profiles but crucially did not investigate Group by Canonicity interactions for *overall* errors,
8
9 choosing instead to focus on different error types. Consequently, there is a need for an
10
11 investigation of profile differences employing GLMs.
12
13
14

The focus of the study

15
16
17
18 This study uses a GLM to analyse data from the Sentence Repetition (SR) paradigm,
19
20 otherwise known as Elicited Imitation. While this is ostensibly a measure of verbatim recall, and
21
22 hence may depend on both STM and WM (Jefferies, Lambon-Ralph, & Baddeley, 2004; Willis &
23
24 Gathercole, 2001), it is also argued that SR involves linguistic representations in long-term
25
26 memory (LTM) (Clay, 1971; Potter & Lombardi, 1998; Slobin & Welsh, 1968). Short-term
27
28 memory (STM) may not have sufficient capacity to support recall of sentences above a certain
29
30 length, and therefore syntactic lexical and semantic representations in LTM are recruited (Potter &
31
32 Lombardi, 1998). Effectively, the sentence is not parroted, but *reconstructed* from activated
33
34 representations in LTM. Numerous studies have demonstrated the involvement of underlying
35
36 syntactic representations. Firstly, Potter and Lombardi (1998) observed structural priming effects
37
38 during SR, and these are widely assumed to involve underlying syntactic representations.
39
40 Secondly, canonicity impacts upon repetition performance even when length is held constant and
41
42 lexical factors are controlled for (Hudgins & Cullinan, 1978; Riches, 2012). Consequently, greater
43
44 errors for non-canonical sentences must be due to structural factors. Finally, repetition of complex
45
46 sentences yields consistent error patterns, e.g. transforming object relatives into subjective
47
48 relatives (Riches et al., 2010), and these cannot be explained without invoking syntactic
49
50 representations.
51
52
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

According to the RECONSTRUCTION HYPOTHESIS, there is a strong overlap between the cognitive mechanisms engaged by SR, and tasks which are regarded as more naturalistic or ecologically valid, e.g. elicitation and forced-choice comprehension tasks. For example, during elicitation tasks using picture prompts, the participant must assemble the sentence from linguistic representations in LTM. According to the reconstruction hypothesis, SR involves essentially the same process, except that representations are primed by the stimulus, so the sentence is not built ‘from scratch.’ Comprehension is also essential to SR as if the sentence is poorly understood, the appropriate representations in LTM will not be activated and the sentence will not be correctly recalled. This was demonstrated by McDade, Simpson and Lamb (1982), who found a strong association between comprehension and recall accuracy for the same stimuli. Overall, most researchers employing SR generally assume that it activates the language system at a deep level. In fact it has been argued that there is ‘general agreement by researchers’ that SR can be used to assess the child’s ‘productive linguistic capacity’ (Bernstien Ratner, 2000, p. 293; as cited in Seeff-Gabriel, Chiat, & Dodd, 2010).

37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

In addition to its cognitive underpinnings, SR offers practical advantages as it is relatively easy to score, given that there is a single target. In contrast, with more open-ended paradigms, e.g. elicitation, there may be more than one correct response. It is beneficial to have a single target because we can reliably quantify the distance between the target and response using an algorithm such as the Levenshtein Distance (LD) (Levenshtein, 1966). An adapted version of this algorithm counts the minimum number of word / morpheme additions, omissions and substitutions required to transform one sentence into another (see Appendix 2 for worked examples – all appendices are in the online supplementary materials). This yields a wide measurement scale with no theoretical upper limit, though in reality the number of errors is

Repetition of complex sentences in SLI

1
2
3 unlikely to exceed the number of words / morphemes in the sentence, corresponding to a null
4
5 response. Such a scale is beneficial because it increases statistical power and provides a sensitive
6
7 measure of performance, which may correspond to the underlying strength of the syntactic
8
9 representation. Additionally, it deals unproblematically with null responses, which are difficult to
10
11 analyse using other paradigms (see Hakansson & Hansson, 2000 for a discussion of this issue).
12
13
14

Aims and hypotheses

15
16
17
18 The study investigated whether children with SLI exhibit a qualitatively different profile
19
20 to language-typical peers on a task involving the production of complex sentences (SR). The study
21
22 incorporated two methodological innovations. Firstly, a GLM was employed to model the
23
24 distribution of the data. Secondly, it used the LD, which increases statistical power and provides a
25
26 sensitive performance metric.
27
28
29

30
31 The main hypothesis was that there would be a significant difference in profiles as
32
33 manifested by a Group by Canonicity interaction. In addition to the LD, word order errors were
34
35 coded to provide a more qualitative measure of performance. Again it was predicted that word
36
37 order errors would also present with a Group by Canonicity interaction.
38
39
40

METHOD

Participants

41
42
43
44
45
46
47 17 children with SLI aged 6;0 to 7;3 were recruited from language units attached to
48
49 mainstream schools in the South East of England. Recruitment letters were sent to Speech and
50
51 Language Therapists, requesting that children meet criteria for SLI, with structural language
52
53 difficulties, English as their main language, and no non-verbal learning difficulties, hearing
54
55 difficulties, autism spectrum disorders, or other known syndrome. No child had been diagnosed
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 with a disorder interfering with intelligibility, e.g. dyspraxia or oromotor difficulties, according to
4
5 a screening questionnaire. Nonverbal abilities were assessed using the Wechsler Preschool and
6
7 Primary Scale of Intelligence core subtests (WPPSI-3: Wechsler, 2002) with all children obtaining
8
9 standard scores greater than or equal to 85. Three assessments were used for assessing structural
10
11 language difficulties; Word Structure (WS) from the CELF (Wiig, Secord, & Semel, 1992), the
12
13 Renfrew Action Picture Task (RAPT: Renfrew, 1997), and the Test of Reception of
14
15 Grammar-Electronic (TROG-E: Bishop, 2005). WS and RAPT assess expressive syntax, with
16
17 both tests designed to elicit specific syntactic structures at both morpheme and sentence level. The
18
19 TROG-E was chosen to assess receptive syntax. This version of the CELF was chosen as it is
20
21 standardized across a wide age range, allowing the same assessment to be used with all children.
22
23 Children were diagnosed with SLI if they fell below -1.2 standard deviations on 2 or more of these
24
25 structural language assessments. In addition to these diagnostic assessments, the British Picture
26
27 Vocabulary Scales (BPVS: Dunn, Whetton, & Burley, 1997), the CELF Recalling Sentences (RS)
28
29 task and the Children's test of Nonword Repetition (CNRep: Gathercole & Baddeley, 1996) were
30
31 also administered.
32
33
34
35
36
37
38
39

40 17 age-matched (AM) and 21 Language-matched (LM) children (age 4;0 to 5;0) were
41
42 recruited from mainstream schools and nurseries via head teachers, with language matching
43
44 accomplished via MLU-in-words (MLUw). Identical instruments were used, with every child
45
46 scoring > 85 on the WPPSI, and no child scoring < -1.2 standard deviations on more than one
47
48 language assessment. Narratives were elicited from the children in order to calculate their
49
50 MLU-in-words (MLUw) for group-matching purposes. The two narratives were the *Bus Story*
51
52 (Renfrew, 1991) and *Frog, Where Are You* (Mayer, 1969), often referred to as the *Frog Story*.
53
54 While the Bus Story involves the experimenter telling the story first, the Frog Story involves the
55
56
57
58
59
60

Repetition of complex sentences in SLI

child building their own narrative from pictures. This narrative-based method is different to the play-based scenario typically used to derive MLU, but has the advantage that it is less influenced by interactional context, and may be more linguistically demanding, eliciting longer and more complex utterances (Hewitt, Hammer, Yont, & Tomblin, 2005; Miles, 2006). This, in turn, may enhance its sensitivity as a language assessment. The children's speech was transcribed using conventions proposed by Miller (1981). Samples contained mean 65.4 utterances (s.d. 20.9) in the SLI group, mean 68.4 (s.d. 14.6) in the AM group, and mean 55.1 (s.d. 16.1) in the LM group.

Table 1 shows psychometrics and significant group differences.

--- INSERT TABLE 1 ABOUT HERE ---

Stimuli

100 sentences were generated according to a 2 (canonicity) x 2 (length) design.

Non-canonical sentences were object relatives and object questions. For relative clauses a mixture of right-branching and centre-embedded clauses were used. Examples are shown in Appendix 1.

Sentences were created in pairs such that for each non-canonical sentence there was a canonical sentence of *exactly the same length* and *employing exactly the same words*. Therefore greater errors on non-canonical sentences cannot be ascribed to lexical and phonological factors. In addition, 2-place and 3-place predicate transitive sentences and passives were used as filler items.

Length ranged from 6 to 12 words (mean 8.2) and was manipulated using filler adjectives and adverbs (see Appendix 1 for examples and frequencies of each construction). All nouns and verbs have a token frequency > 10 words per million on either the British National Corpus (*British National Corpus, version 2, 2002*), or the CELEX database; spoken and written (Burnage, 1990).

All stimuli were spoken by a native female speaker of English with a local dialect, and recorded in a sound-proof booth. Sentences were grouped into 8 blocks of 20 and pseudorandomised so that no

Repetition of complex sentences in SLI

1
2
3 two consecutive sentences had the same type, length and canonicity characteristics. Sentence
4
5 length gradually increased throughout the block, as in piloting, this facilitated performance on the
6
7 longer sentences, a method also adopted by standardised SR assessments, e.g. the CELF.
8
9

10
11 17 linguistically-informed colleagues rated the plausibility of sentence pairs, e.g. *the*
12
13 *monkey chased the pig*, and *the pig chased the monkey*, and sentences were only chosen if there
14
15 was a small discrepancy between these (a maximum of 3 points on a rating scale of 7) indicating
16
17 that both propositions were more or less equally probable. These simple transitive sentences were
18
19 then used to create the stimuli, e.g. *which pig is the monkey chasing?* This process ensured that the
20
21 sentences were highly reversible, i.e. changing the order of the arguments (*the monkey chased the*
22
23 *pig* versus *the pig chased the monkey*) does not greatly affect the plausibility of the sentence.
24
25 Reversibility is an important factor to control for as it may affect the likelihood that a child will
26
27 make word order errors. For example, a child would be very unlikely to reverse the word order in a
28
29 non-reversible sentence such as *the man drew the picture*.
30
31
32
33

*Procedure*Administration

34
35
36
37
38
39 The SR task was demonstrated with a cuddly toy parrot and a story book called the
40
41 Gossipy Parrot (Roddie & Terry, 2003). The experimenter read the story to the child, and at
42
43 various stages the parrot commented on the story. This was achieved wirelessly via a Kensington
44
45 conference pointer hidden inside the toy. The experimenter pretended not to understand the parrot,
46
47 so the child had to help him by repeating what the parrot had said. The parrot was also used for the
48
49 SR task itself which was run on a laptop computer. The experimenter said 'Now the parrot is going
50
51 to say some more sentences. I don't understand parrots so you have to tell me exactly what the
52
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 parrot says'. The child was then presented with a 5 x 4 grid, with a coloured band to show
4
5 half-way. As the child heard each sentence a number appeared in the grid. This motivated the child
6
7 by showing how many sentences remained. At the end, a 'reward' screen appeared with a picture
8
9 of people clapping accompanied by applause. All sentences were heard via headphones
10
11 (Sennheiser PC156), and responses were recorded to the computer via the mouthpiece. The
12
13 experiment was run using DMDX experimental software (Forster & Forster, 2003).
14
15
16
17

18
19 Assessments were conducted during 3 visits per child. Each visit consisted of two 30-40
20
21 minute sessions separated by a break. Sentence repetition blocks were administered in one of 4
22
23 pseudorandomised orders, with orders evenly distributed within groups.
24
25

Coding

26
27
28 Quantitative errors were derived using the LD. For the purpose of this analysis, each
29
30 morpheme was represented as a separate unit. Therefore, this measure will be described as the
31
32 Levenshtein Distance in morphemes (LDm). By coding sentences in terms of morphemes,
33
34 omissions of affixes will be counted by the algorithm (see Appendix 2A for demonstrations). In
35
36 addition, a more qualitative measure of structural changes was devised. Structure was deemed to
37
38 have changed if any of the arguments or the main verbs changed position or syntactic function, e.g.
39
40 OSV became SVO (e.g. *which pig is the cheeky little monkey rescuing* → *which pig is the rescuing*
41
42 *the cheeky little monkey*) or SVO became OSV. A second rater coded responses for 2 AM children,
43
44 3 LM children and 4 children with SLI, corresponding to 15% of the observations. Ratings were
45
46 identical for 96% of responses.
47
48
49
50
51

Elicitation

1 Repetition of complex sentences in SLI

2
3 A central argument of the study is that SR can be used to assess representations in LTM.
4
5 To verify this claim, an elicitation task was conducted which depended on structural priming. It
6
7 has been widely argued that where there is no lexical overlap structural priming reflects syntactic
8
9 representations in long-term memory (e.g. Pickering & Ferreira, 2008). During the elicitation task,
10
11 the experimenter described one picture using the target structure, and the child was encouraged to
12
13 describe a different picture, e.g. *EXPERIMENTER: This is the bread which the woman baked and*
14
15 *this is the soup... CHILD: which the boy made.* While this is not a standard priming paradigm,
16
17 nonetheless the children are primed to reproduce the structure in the first clause. Importantly, all
18
19 responses required a change of both verb and noun, thereby pre-empting verbatim recall, and STM
20
21 demands were minimal, as completion fragments number no more than five words. This entailed
22
23 producing the head of the relative clause which precluded the possibility of reversing arguments,
24
25 e.g. *this is the boy who made the soup.* The assessment contained 2 warmup items, 2 subject
26
27 relatives and 2 object relatives. An attempt was also made to prime questions, but this proved
28
29 difficult. A scoring protocol was devised to reflect the main syntactic components of each structure
30
31 (see Appendix 2B), with scores ranging from zero to one.
32
33
34
35
36
37
38
39

40 ANALYSIS

41
42
43 *Analysis of the validity of the dependent variables*

44
45 Because the LDm is a novel metric an analysis of validity was conducted (Table 2).
46
47 Measures were also obtained for a qualitative measure; the number of responses per child
48
49 containing word order errors, e.g. OSV → SVO. Analyses investigated construct validity, i.e.
50
51 whether the assessment was a good reflection of performance on complex sentences; the
52
53 theoretical area of interest, and concurrent validity, i.e. whether they are associated with other
54
55 measures which have strong empirical support as a clinical marker of SLI. Construct validity was
56
57
58
59
60

Repetition of complex sentences in SLI

determined by (a) investigating whether the dependent variable was influenced by canonicity, and (b) examining the association between the dependent variable and performance on the elicitation task, which demonstrates better ‘face validity’. In other words, elicitation is subjectively a more obvious measure of language production, given that children must actually create a sentence, than SR as children merely need to repeat what the experimenter has said. For the latter, only repetitions of relative clauses were included in the analyses, to ensure that the elicitation and repetition paradigms involved identical structures. With regard to concurrent validity, the decision was made to adopt the CNRep as a dependent variable as it is a reliable clinical marker of SLI (Conti-Ramsden, Botting, & Faragher, 2001). The TROG was also used as a dependent variable. Though not regarded as a clinical marker, it is nonetheless a widely used, well-standardised and well-validated assessment of language abilities. The LDm demonstrated good validity across all measures except the CNRep, thereby demonstrating sensitivity to both syntactic knowledge / abilities and degree of language impairment. By contrast, the qualitative measure presented with good validity on the first measure only.

--- INSERT TABLE 2 ABOUT HERE ---

Data diagnostics and choice of statistical model

The distribution of the error data (LDm) was investigated to determine the choice of regression model. The standard method for modelling count data, e.g. numbers of errors, is a poisson regression. However, the raw data were strongly rightward skewed (see Appendix 3). This distribution is characteristic for count data where there are high rates of zero values (i.e. sentences where no error was made) resulting in ‘overdispersion’ (where the variance is greater than the mean). There are a couple of methods for dealing with overdispersion. Firstly, an extra parameter accounting for the relationship between the mean and variance may be added. This type of

Repetition of complex sentences in SLI

1
2
3 regression is called a negative binomial regression. In addition, zero inflation (or ‘zero truncation’)
4
5 may be applied. This combines a poisson model, with a logit model to account for excess zeroes.
6
7
8 The appendix shows statistics of model fit for a variety of models using two modules in R (R
9
10 Development Core Team, 2014); lme4 (Bates, Maechler, Bolker, & Walker, 2014), and
11
12 glmmADMB (Skaug, Fournier, Nielsen, Magnusson, & Bolker, 2011). It can be seen that the
13
14 best-fitting model was the zero-inflated negative binomial model, which significantly improves on
15
16 the next best model.
17
18

Analysis of errors (LDm)

19
20
21
22
23 Descriptives are shown in Table 3 and Figure 1. Differences in raw errors as a function
24
25 of canonicity clearly vary across groups, with a larger difference observed in the SLI group. This is
26
27 consistent with the idea that children with SLI find non-canonical sentences especially difficult.
28
29 However, when data are presented as ratios a different picture emerges. The ratio shows the
30
31 number of errors in non-canonical sentences per error in canonical sentences. For example, the
32
33 children with SLI made 1.22 errors in non-canonical sentences for each error in canonical
34
35 sentences. The ratio is a mirror image of the difference data, with the SLI groups exhibiting the
36
37 smallest ratios, and the LM children exhibiting the largest ratios.
38
39
40
41
42

43 --- INSERT TABLE 3 AND FIGURE 1 ABOUT HERE ---
44
45

46 To begin with, a traditional by-subjects ANOVA was conducted with Group and
47
48 Canonicity as the independent variable, and the LDm as the dependent variable. The LDm was
49
50 divided by the number of morphemes in the sentence to give a rate variable (errors per morpheme).
51
52 This ensures that the ANOVA is consistent with the GLMs (below) which also modelled the errors
53
54 as rates. The mean rate was then calculated for each participant by Canonicity combination. There
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 was a significant effect of Group ($F(2, 52) = 51.3, p < 0.001^{***}, \eta^2_p = 0.969$), a significant effect of
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

was a significant effect of Group ($F(2, 52) = 51.3, p < 0.001^{***}, \eta^2_p = 0.969$), a significant effect of
Canonicity ($F(1, 52) = 95.5, p < 0.001^{***}, \eta^2_p = 0.647$) and a significant interaction ($F(2, 52) =$
 $4.07, p = 0.023^*, \eta^2_p = 0.135$). Planned contrasts found a significant effect for the SLI versus AM
comparison ($F(1)=7.76, p < 0.007^{**}$), but the SLI versus LM contrast just missed significance
($F(1)= 3.92, p = 0.053$).

A mixed effects negative binomial regression with zero inflation was conducted. Like
the ANOVA this modelled the effects of Group, Canonicity and the Interaction term. Raw errors
per sentence were entered as the dependent variable. In order to turn these into rates, sentence
length-in-morphemes was set as the ‘exposure’, i.e. the size of the unit within which the errors
occurred (glmmADMB only provides an ‘offset’ option, and to transform this into the exposure we
take the square root of the rate measure; the number of morphemes). Group was treatment coded,
with the SLI group specified as the reference group. Analyses adopted a maximal random effects
structure, with participant and item entered as intercepts, and by-participant slopes for canonicity
(Barr et al., 2013). The association between participant (intercept) and Canonicity (slope) was low
($r = -0.14$), and therefore the latter was retained (Baayen et al., 2008). Though the glmmADMB
analysis reports significance (see Table 4) an alternative procedure using likelihood ratio tests was
also conducted as it is thought to provide better estimates (Barr et al., 2013). Coefficients have
been reported as Incidence Rate Ratios (IRR) to aid interpretation. The IRR for the SLI versus AM
contrast was 0.23, signifying that each child in the AM group made 0.23 errors for each error made
by a child in the SLI group. Overall there was a significant effect of Group, the effect of Canonicity
just missed significance, and the Group by Canonicity interaction did not approach significance.

1 Repetition of complex sentences in SLI

2
3 *Analysis of qualitative errors*

4
5 Descriptives are shown in Table 5. Most errors involve both word order changes and the
6 swapping of thematic roles, e.g. there's the cat that the dog chased → there's the cat that chased
7 the dog. Here an object relative (within a presentational cleft) is changed into a subject relative, but
8 the nouns have remained in the same positions, thereby swapping thematic roles (and also
9 syntactic functions). As this error type was by far the most common it was selected for the
10 subsequent analysis. The bottom rows in the table show this error type summarised by condition.
11 The children with SLI appear to be particularly sensitive to the structure of the sentence,
12 demonstrating a strong tendency to transform canonical into non-canonical sentences.
13
14
15
16
17
18
19
20
21
22
23
24

25 --- INSERT TABLE 5 ABOUT HERE ---
26
27

28 The data were subsequently analysed using a mixed effects logistic regression to model
29 dichotomous data (Jaeger, 2008). The independent variables were Group, Canonicity and their
30 interaction, and the dependent variable was coded 1 if the child made the error type described
31 above, and 0 if such an error was not evident. Details of the model-fitting procedure are shown
32 below the table. There was a significant main effect of Canonicity, but no significant effect of
33 Group. The Group by Canonicity interaction demonstrated a trend towards significance.
34
35
36
37
38
39
40
41
42
43

44 --- INSERT TABLE 6 ABOUT HERE ---
45

46
47 *Investigation of individual differences in the SLI group*

48
49 Histograms were plotted to investigate individual differences in both the LDm and
50 above error type. While LDm performance was relatively homogenous, with few outlying children
51 in any of the groups, histograms identified 3 children with SLI who make high rates of such errors
52 (see Figure 2). Scores on other language assessments were inspected to find out if these children
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 belong to a Grammatical or Syntactic SLI subgroup (Friedmann & Novogrodsky, 2011; van der
4
5 Lely, 2005). Scores are shown in Table 7. The possibility of a subgroup was partially corroborated
6
7
8 for children 3 and 11. They were both outliers (performance + or - 1 standard deviation) on the
9
10 ratio measure, while 3 was also an outlier on the difference measures. Both of these measures are
11
12 sensitive to extreme difficulties with non-canonical structures. However, there was no evidence
13
14 that these children were outliers on more general linguistic measures.
15
16

17
18 --- INSERT TABLE 7 AND FIGURE 2 ABOUT HERE ---
19
20

DISCUSSION

21
22
23
24
25 Complex sentence profiles in SLI were investigated using SR. The ANOVA identified a
26
27 significant interaction between Group and Canonicity when the children with SLI were compared
28
29 to the AM group. Inspection of the means (Table 3, Figure 1) indicates a larger effect of
30
31 Canonicity in the SLI group, suggesting that this interaction is driven by specific difficulties with
32
33 non-canonical sentences in this group. A group by canonicity interaction for the SLI versus AM
34
35 comparison is consistent with a number of previous studies of complex sentence profiles
36
37 (Friedmann & Novogrodsky, 2011; Jensen de López et al., 2012; Riches et al., 2010). However,
38
39 the interaction term for the SLI versus LM contrast just missed significance ($p = 0.053$). In contrast
40
41 to the ANOVA, a negative binomial model regression with zero inflation did not identify any
42
43 differences in profiles. This finding conflicts with the majority of the research literature. A
44
45 qualitative measure of word order errors identified a trend towards a significant interaction
46
47 between Group and Canonicity ($p = 0.078$), such that children with SLI tended to transform
48
49 non-canonical into canonical sentences. This effect was driven by a subgroup of three children
50
51 who made elevated rates of such errors.
52
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 An important finding of the study is that the significance of the interaction term is
4
5 dependent on our choice of statistical model. While the ANOVA identified a significant
6
7 interaction (SLI versus AM) none was observed for the negative binomial regression with zero
8
9 inflation. Moreover, differences in p-values were large. This has strong theoretical implications as
10
11 interpretation of the interaction term determines whether we view SLI as a disorder characterised
12
13 by severe deficits in particular subcomponents of the language system, or systems subserving
14
15 language. There are two statistical arguments in favour of the GLM, as outlined in the literature
16
17 review. Firstly, the best-fitting regression (negative binomial with zero inflation) allowed us to
18
19 model changes in variance as the data approached the extremes. GLMs are more effective than
20
21 ANOVAs in this regard (Jaeger, 2008). Secondly, the GLM incorporated a fully-specified random
22
23 effects structure, modelling the effect of both items and participants. Both of these factors allow
24
25 for a more reliable investigation of interaction effects (Barr et al., 2013; Clark, 1973; Jaeger,
26
27 2008).
28
29
30
31
32
33
34

35 One way to conceptualise the difference between the two analyses is to think in terms of
36
37 additive and multiplicative models. According to the GLM, the Incidence Rate Ratio for the
38
39 Canonicity term is 1.26. This means that to obtain error rates for non-canonical sentences, we must
40
41 multiply error rates for canonical sentences by 1.26. This multiplicative relationship stems from
42
43 the use of a log link function, as when logs are added the underlying bases are multiplied, e.g. $e^{\log(x)}$
44
45 $+\log(y) = x * y$ (e = natural logarithm). By contrast, ANOVAs on untransformed data use an additive
46
47 model based on absolute differences between conditions. This difference matters when
48
49 investigating interaction effects. For example, glancing at Figure 1, absolute differences between
50
51 conditions are greater in the SLI group, which drives the significant Group by Canonicity
52
53 interaction in the ANOVA on untransformed data (SLI versus AM contrast only). However, the
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 relative heights of the bars vary little across the groups, with the mean for non-canonical sentences
4
5 approximately 26% higher than the mean for canonical sentences. This leads to an absence of
6
7 profile differences when analysed using the GLM.
8
9

10
11 The use of a multiplicative model is justified by the statistical tests of model fit.
12
13 However, it is also strongly desirable to develop a psycholinguistic model which operates in a
14
15 multiplicative fashion. In fact, most formal accounts of language processing, i.e. those adopting
16
17 algebraic notation, posit multiplicative relationships. For example, Gibson (1998, pg. 16) argues
18
19 that the speed of integration at the trace is equal to; $\text{Constant} * (\text{Energy Resources} / \text{Available}$
20
21 $\text{Memory Resources})$. If we imagine that the Constant varies across groups, and Energy Resources
22
23 vary as a function of Canonicity, then we have a theoretical framework consistent with our
24
25 multiplicative statistical model. Just and Carpenter's (1992) CC-reader also posits multiplicative
26
27 relationships. For example, when activation flows from one 'element' to another it is multiplied by
28
29 a constant (pg. 15). This mirrors the basic design of artificial neural networks where synaptic
30
31 weights multiply the activation by a given value.
32
33
34
35
36
37

38 In addition to these theoretical arguments, there is also an empirical measure which, in
39
40 studies of children with SLI, has consistently demonstrated multiplicative properties; speed of
41
42 processing. Children with SLI are about a third slower than control children across a range of
43
44 linguistic and non-linguistic tasks (Kail, 1994). In other words, we multiply the RTs of typically
45
46 developing children by 1.33 to obtain the SLI data. Unfortunately, there is currently little evidence
47
48 identifying speed of processing as a primary determinant of language difficulties (Leonard et al.,
49
50 2007). Nonetheless, studies of processing speed in SLI lend validity to formal models positing
51
52 multiplicative relationships.
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

If we accept the statistical and theoretical arguments outlined above we must conclude that profiles, as manifested by an interaction effect, do *not* vary across groups. Moreover the discrepancies between the results for the GLM and the ANOVA support the claim that the latter can spuriously inflate the significance of interaction terms (Jaeger, 2008). Given these limitations we should be extremely cautious when using interaction terms, combined with an inspection of mean scores, to infer a qualitative difference in performance characterised by a localised deficit. In addition, if there are no genuine differences in profiles, we ought to conceptualise children with SLI as ‘low-language’ children, i.e. children at the tail end of the normal distribution, as opposed to a distinct diagnostic category. While this viewpoint differs from most experimental studies it is nonetheless consistent with large scale epidemiological studies. For example, in a study of 1,529 children, Tomblin and Zhang (2006) found that a single factor explained performance on range of different language assessments measuring expressive and receptive vocabulary, and expressive and receptive sentence-level syntax. There also exist theoretical models which account for a range of linguistic difficulties in SLI using a single language-related factor, thereby undermining the claim for localised difficulties. For example, it has been suggested that lexical and morphological development are closely synchronised, with lexical learning providing the raw materials for acquiring morphemes (Conti-Ramsden & Jones, 1997).

Though calling into question theories of profile differences, the data did support the claim that children with SLI have difficulties with thematic role assignment (Friedmann & Novogrodsky, 2007). While changing object relatives into subject relatives, they failed to switch Noun Phrases in order to preserve meaning, e.g. there’s the cat that the dog chased → there’s the cat that chased the dog, and were therefore oblivious to thematic role changes. It is also true that thematic role errors tended to occur in object relatives, which at first glance supports the argument

Repetition of complex sentences in SLI

1
2
3 that thematic role assignment is particularly difficult in the context of non-canonical sentences.
4
5 However, thematic role errors were only observable precisely *because* the non-canonical
6
7 sentences resulted in changes to word order. As canonical sentences rarely elicited such word
8
9 order errors, it is difficult to determine whether these posed difficulties with thematic role
10
11 assignment. Overall, the data indicate difficulties with thematic role assignment, but arguably, the
12
13 paradigm is not well suited for investigating whether these are subject to structural constraints.
14
15
16

Limitations

17
18
19
20 From a statistical perspective, there are a number of potential difficulties with the study.
21
22 The claim of no differences in profiles is based on a null result which may reflect limited statistical
23
24 power. It could be argued that the confidence intervals for the crucial interaction term (Table 4)
25
26 were relatively narrow suggesting good generalisation to different populations and / or items.
27
28 Unfortunately, there are no objective criteria for determining ‘acceptable’ intervals. Nonetheless,
29
30 even if the study were underpowered, the contrasting findings for the ANOVA and GLM support
31
32 the key claim that one’s choice of model may critically affect the interaction term. Another
33
34 statistical issue is the strong rightward skew in the data, which required a complex analysis. It
35
36 could be argued that this kind of data is atypical, and therefore should not form the basis for more
37
38 general claims related to quantitative methods. However, given that near-ceiling, or near-floor
39
40 performance is widely observed in the research literature (e.g. Deevy & Leonard, 2004;
41
42 Novogrodsky & Friedmann, 2006), it is likely that the distribution observed in the current study is
43
44 relatively common. A further statistical issue is the nature of the dependent variable. The LDm is
45
46 only quasi-count, in that it calculates many different types of errors (addition, substitution,
47
48 omission), and the minimum distance can involve more than one set of operations. However, count
49
50 models (e.g. negative binomial) clearly fit the distribution of the data better than non-count models
51
52
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 (Gaussian), suggesting that treating the LDm as a count variable is statistically justifiable. Finally,
4
5 it should be noted that the random effects for the logistic regression needed to be simplified for the
6
7 model to converge, which may reflect limited statistical power. Such simplification is likely to
8
9 increase p-values, and therefore impact on moderately significant effects. As there were no
10
11 moderately significant terms it is debatable whether simplification of random effects greatly
12
13 impacts on the findings.
14
15
16
17

18
19 Moving on to theoretical considerations, it could be argued that the study did not find
20
21 different profiles because it failed to recruit children with genuine syntactic difficulties. In support
22
23 of this there was clear heterogeneity with regard to qualitative error profiles, with a group of three
24
25 children with SLI exhibiting a very strong tendency to transform non-canonical into canonical
26
27 sentences. This is consistent with claims for a subgroup of language-impaired children with
28
29 particularly severe syntactic difficulties (Friedmann & Novogrodsky, 2011; van der Lely, 2005).
30
31 Analysis of other measures partially supported this interpretation with two children exhibiting a
32
33 particularly high error rate on non-canonical sentences. This heterogeneity suggests that the
34
35 screening measures may not have succeeded in identifying children with genuine structural
36
37 language difficulties, and this would account for the failure to identify an atypical profile in the
38
39 group as a whole. While this remains a possibility it should be noted that the language screening
40
41 measures did assess structural abilities, e.g. syntactic morphemes (CELF-WS) and complex
42
43 sentences (TROG). Moreover, setting aside debates on screening protocols, the discrepant findings
44
45 for the ANOVA and GLM support the key claim that inferences based on interaction terms are
46
47 problematic. Overall, though the study critiques claims of a localised syntactic deficit in SLI, it
48
49 nonetheless raises the possibility of a relatively rare SLI subgroup exhibiting localised syntactic
50
51
52
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 difficulties. To resolve this issue it is clearly important that studies investigate individual as well as
4
5 group profiles.
6
7

8
9 Another potentially problematic aspect of the study is the relatively strong performance
10
11 of the LM children on a number of language assessments, though crucially, not the MLUw, which
12
13 was adopted as the matching variable. This complicates the language-matching process.
14
15 Nonetheless, it is unlikely that a better-matched LM group would have impacted on the main
16
17 finding of the study; the lack of a significant profile difference according to the GLM.
18
19

20
21 A final extremely important point to make is that analyses of cross-sectional data are
22
23 limited as they do not reveal developmental trajectories. Highly powered longitudinal studies at
24
25 the macro-level, i.e. across linguistic domains, demonstrate that children with SLI are more
26
27 delayed in their performance on morphosyntactic assessments, than they are on vocabulary
28
29 assessments (Rice, 2013). These growth curves indicate that certain subdomains, e.g.
30
31 morphosyntax, may be more severely affected/delayed than others. In addition, the current study
32
33 identified a further asynchrony, with the SR performance of the children with SLI lagging behind
34
35 those of developmental controls. In conclusion, while inferences based on cross-sectional data are
36
37 problematic, existing evidence for qualitative differences from longitudinal data is harder to
38
39 dismiss. Consequently, longitudinal datasets may provide a better way of addressing the
40
41 delay-versus-difference debate.
42
43
44
45
46
47

Future directions

48
49 It is widely argued that SLI is characterised by qualitatively unusual profiles. However,
50
51 the current analysis suggests that existing methods to determine profile differences in
52
53 cross-sectional data may be unreliable, and GLMs may offer a better alternative. If the claims are
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 correct then the move away from ANOVAs towards GLMs, currently popular in the research
4
5 literature, may result in a weakening of the claim for distinct syntactic profiles, and towards a
6
7 conception of SLI as 'low-language' children. It would also be interesting to find out whether
8
9 GLM analyses of already-published data further support the claim that children with SLI are
10
11 qualitatively distinct. As a caveat, the analysis does not address other evidence in favour of SLI as
12
13 a distinct category, including the existence of developmental asynchronies (Rice, 2013), and
14
15 idiosyncratic error types, such as difficulties assigning thematic roles (Friedmann & Novogrodsky,
16
17 2007).

18
19
20
21
22
23 While there is strong evidence to support the use of SR as measure of linguistic
24
25 competence (Potter & Lombardi, 1998), it is a relatively artificial task and is likely to tap into other
26
27 mechanisms such as Phonological STM (Riches, 2012). Therefore, it is recommended that GLMs
28
29 be used to analyse profiles in more ecologically valid comprehension and elicitation tasks. GLMs
30
31 could also be extended to investigate other types of profiles, for example profiles within linguistic
32
33 subdomains, for example, regular versus irregular verb morphology as investigated by van der
34
35 Lely & Ullman (2001).
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Acknowledgements

50
51
52
53 I would like to thank the British Academy for funding this project (Award Number
54
55 PDF/2007/460), and colleagues at the Universities of Reading and Newcastle for their support and
56
57
58
59
60

Repetition of complex sentences in SLI

1
2
3 advice. In particular I would like to thank Thomas King, David Howard, Christos Salis and Nils
4
5 Braakmann for their statistical input. Thanks also to Kerry Davis and Christos Pliatsikas for help
6
7
8 with data collection and coding and the three anonymous reviewers. Finally, a massive thank you
9
10 for those teachers, speech and language therapists, and above all the children who participated.
11
12
13

14
15
16
17 Adani, F., Forgiarini, M., Guasti, M. T., & van der Lely, H. K. J. (2014). Number dissimilarities
18
19 facilitate the comprehension of relative clauses in children with (Grammatical) Specific
20
21 Language Impairment. *Journal of Child Language*, 41(4), 811 – 841.
22
23

24
25 Arnon, I. (2010). Rethinking child difficulty: The effect of NP type on children's processing of
26
27 relative clauses in Hebrew. *Journal of Child Language*, 37(01), 27–57.
28
29

30
31 Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed
32
33 random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
34
35

36
37 Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for
38
39 confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*,
40
41 68(3), 255–278. <http://doi.org/10.1016/j.jml.2012.11.001>
42
43

44
45 Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using*
46
47 *Eigen and S4*. Retrieved from <http://CRAN.R-project.org/package=lme4>
48

49
50 Bernstien Ratner, N. (2000). Elicited Imitation and Other Methods for the Analysis of Trade-Offs
51
52 Between Speech and Language Skills in Children. In L. Menn (Ed.), *Methods for studying*
53
54 *language production*. (pp. 291 – 312). Lawrence Erlbaum Assoc Inc.
55
56
57
58
59
60

Repetition of complex sentences in SLI

- 1
2
3 Bishop, D. V. M. (2005). *Test of Reception of Grammar - Electronic*. London: Harcourt
4
5 Assessment.
6
7
8
9 Booth, J. R., MacWhinney, B., & Harasaki, Y. (2000). Developmental differences in visual and
10
11 auditory processing of complex sentences. *Child Development*, 71(4), 981 – 1003.
12
13
14 *British National Corpus, version 2*. (2002). Oxford: BNC Consortium.
15
16
17
18 Burnage, G. (1990). *CELEX: A Guide for Users*. Nijmegen: CELEX.
19
20
21 Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in
22
23 psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12(4), 335–359.
24
25
26
27 Clay, M. M. (1971). Sentence Repetition: Elicited Imitation of a Controlled Set of Syntactic
28
29 Structures by Four Language Groups. *Monographs of the Society for Research in Child*
30
31 *Development*, 36(3), 1–85.
32
33
34 Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific
35
36 language impairment (SLI). *The Journal of Child Psychology and Psychiatry and Allied*
37
38 *Disciplines*, 42(6), 741–748.
39
40
41
42 Conti-Ramsden, G., & Jones, M. (1997). Verb use in Specific Language Impairment. *Journal of*
43
44 *Speech Language and Hearing Research*, 40, 1298 – 1313.
45
46
47
48 Deevy, P., & Leonard, L. B. (2004). The comprehension of wh-questions in children with Specific
49
50 Language Impairment. *Journal of Speech Language and Hearing Research*, 47(4), 802 –
51
52 815.
53
54
55
56
57
58
59
60

Repetition of complex sentences in SLI

- 1
2
3 de Villiers, J. G. (2003). Defining SLI: A linguistic perspective. In Y. Levy & J. Schaeffer (Eds.),
4
5 *Language competence across populations: Toward a definition of specific language*
6
7 *impairment* (pp. 425–447). Mahwah, NJ: Laurence Erlbaum Associates.
- 8
9
10
11 Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and*
12
13 *Language*, 59(4), 447–456. <http://doi.org/10.1016/j.jml.2007.11.004>
- 14
15
16
17 Dollaghan, C. A. (2004). Taxometric analyses of specific language impairment in 3-and
18
19 4-year-old children. *Journal of Speech, Language, and Hearing Research*, 47(2), 464 –
20
21 475.
- 22
23
24
25 Dollaghan, C. A. (2011). Taxometric Analyses of Specific Language Impairment in 6-Year-Old
26
27 Children. *Journal of Speech, Language, and Hearing Research*, 54, 1361 – 1371.
28
29 [http://doi.org/10.1044/1092-4388\(2011/10-0187\)](http://doi.org/10.1044/1092-4388(2011/10-0187))
- 30
31
32
33 Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale*. Windsor,
34
35 England: NFER - Nelson.
- 36
37
38
39 Epstein, B., Hestvik, A., Shafer, V. L., & Schwartz, R. G. (2013). ERPs reveal atypical processing
40
41 of subject versus object *Wh* -questions in children with specific language impairment:
42
43 Atypical wh-question processing in SLI. *International Journal of Language &*
44
45 *Communication Disorders*, 48(4), 351–365. <http://doi.org/10.1111/1460-6984.12009>
- 46
47
48
49 Field, A. (2000). *Discovering statistics using SPSS for Windows*. London: Sage.
- 50
51
52 Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond
53
54 accuracy. *Behavior Research Methods Instruments and Computers*, 35(1), 116–124.
55
56
57
58
59
60

Repetition of complex sentences in SLI

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Friedmann, N., & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain and Language*, *101*(1), 50–63.
<http://doi.org/10.1016/j.bandl.2006.09.006>
- Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand?: The comprehension of Wh questions in three subtypes of SLI. *Lingua*, *121*, 367 – 382.
- Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of non-word repetition*. Hove, UK: The Psychology Press.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1 – 76.
- Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, *2*(7), 262–268.
- Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1411.
- Hakansson, G., & Hansson, K. (2000). Comprehension and production of relative clauses: a comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, *27*, 313–333.
- Hewitt, L. E., Hammer, C. S., Yont, K. M., & Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders*, *38*(3), 197–213.

Repetition of complex sentences in SLI

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Howell, D. C. (2013). *Statistical methods for psychology*. Belmont, CA: Wadsworth Cengage Learning.
- Hudgins, J. C., & Cullinan, W. L. (1978). The effect of sentence structure on sentence elicited imitation responses. *Journal of Speech and Hearing Research, 21*, 809 – 819.
- Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language, 59*(4), 434–446.
<http://doi.org/10.1016/j.jml.2007.11.007>
- Jefferies, E., Lambon-Ralph, M., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language, 51*(4), 623 – 642.
- Jensen de López, K., Olsen, L. S., & Chondrogianni, V. (2012). Annoying Danish relatives: Comprehension and production of relative clauses by Danish children with and without SLI. *Journal of Child Language, (FirstView Articles)*.
- Just, A. J., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*(1), 122 – 149.
- Kail, R. V. (1994). A method of studying the generalised slowing hypothesis in children with Specific Language Impairment. *Journal of Speech and Hearing Research, 37*, 418 – 421.
- Leonard, L. B. (2014). *Children with Specific Language Impairment*. Cambridge, MA: MIT Press.
- Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50*(2), 408–428.

Repetition of complex sentences in SLI

- 1
2
3 Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals.
4
5 *Soviet Physics - Doklady*, 10, 707 – 710.
6
7
8
9 Mayer, M. (1969). *Frog, Where Are You?* New York, NY: Dial.
10
11
12 McDade, H. L., Simpson, M. A., & Lamb, D. E. (1982). The use of elicited imitation as a measure
13
14 of expressive grammar: a question of validity. *Journal of Speech and Hearing Disorders*,
15
16 47, 19 – 24.
17
18
19
20 Miles, S. (2006). Sampling context affects MLU in the language of adolescents with Down
21
22 syndrome. *Journal of Speech, Language, and Hearing Research*, 49(2), 325–337.
23
24
25
26 Miller, J. F. (1981). *Assessing Language Production in Children: experimental procedures*.
27
28 Baltimore: University Park Press.
29
30
31 Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working Memory and Specific
32
33 Language Impairment: An Update on the Relation and Perspectives on Assessment and
34
35 Treatment. *American Journal of Speech-Language Pathology*, 19(1), 78.
36
37
38
39 Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A
40
41 window to the nature of the impairment. *International Journal of Speech-Language*
42
43 *Pathology*, 8(4), 364–375. <http://doi.org/10.1080/14417040600919496>
44
45
46
47 Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological*
48
49 *Bulletin*, 134(3), 427 – 459.
50
51
52
53 Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal*
54
55 *of Memory and Language*, 38(3), 265 – 282.
56
57
58
59
60

Repetition of complex sentences in SLI

- 1
2
3 Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed
4
5 random effects and with binomial data. *Journal of Memory and Language*, 59(4), 413–425.
6
7
8
9 R Development Core Team. (2014). *R: A Language Environment for Statistical Computing*.
10
11 Vienna, Austria: R Foundation for Statistical Computing.
12
13
14 Reilly, S., Tomblin, J. B., Law, J., McKean, C., Mensah, K., Morgan, A., ... Wake, M. (2014).
15
16 Specific Language Impairment: a convenient label for whom? *International Journal of*
17
18 *Language and Communication Disorders*.
19
20
21
22 Renfrew, C. (1991). *The Bus Story: a test of continuous speech* (2nd ed.). Oxford: Published by
23
24 author.
25
26
27
28 Renfrew, C. (1997). *Action Picture Test*. Oxford: Speechmark.
29
30
31 Rice, M. L. (2013). Language growth and genetics of specific language impairment*.
32
33 *International Journal of Speech-Language Pathology*, 15(3), 223–233.
34
35 <http://doi.org/10.3109/17549507.2013.783113>
36
37
38
39 Rice, M. L., Wexler, K., & Cleave, P. (1995). Specific Language Impairment as a period of
40
41 extended optional infinitive. *Journal of Speech and Hearing Research*, 38, 850 – 863.
42
43
44
45 Riches, N. G. (2012). Sentence repetition in children with specific language impairment: an
46
47 investigation of underlying mechanisms. *International Journal of Language &*
48
49 *Communication Disorders*, 47(5), 499–510.
50
51 <http://doi.org/10.1111/j.1460-6984.2012.00158.x>
52
53
54
55 Riches, N. G., Loucas, T., Charman, T., Simonoff, E., & Baird, G. (2010). Sentence Repetition in
56
57 Adolescents with Specific Language Impairments and Autism: an Investigation of
58
59

1 Repetition of complex sentences in SLI

2
3 Complex Syntax. *International Journal of Language and Communication Disorders*,
4
5 45(1), 47 – 60. <http://doi.org/10.3109/13682820802647676>
6
7

8
9 Rizzi, L. (1990). *Relativized minimality*. The MIT Press.

10
11
12 Roddie, S., & Terry, M. (2003). *The Gossipy Parrot*. Bloomsbury.

13
14
15 Seeff-Gabriel, B., Chiat, S., & Dodd, B. (2010). Sentence imitation as a tool in identifying
16
17 expressive morphosyntactic difficulties in children with severe speech difficulties.
18
19 *International Journal of Language & Communication Disorders*, 45(6), 691–702.
20
21 <http://doi.org/10.3109/13682820903509432>
22
23

24
25
26 Skaug, H., Fournier, D., Nielsen, A., Magnusson, A., & Bolker, B. (2011). glmmADMB:
27
28 generalized linear mixed models using AD Model Builder. *R Package Version 0.6, 5*, r143.
29
30

31
32 Slobin, D. I., & Welsh, C. A. (1968). Elicited imitation as a research tool in developmental
33
34 psycholinguistics. *Working Papers of the Language Behavior Research Laboratory*,
35
36 *University of California, Berkeley*, 10.
37
38

39
40 Stavrakaki, S. (2001). Comprehension of reversible relative clauses in Specifically Language
41
42 Impaired and normally developing Greek children. *Brain and Language*, 77(3), 419 – 431.
43
44

45
46 Tomblin, J. B., Records, N. L., Buckwalter, P., & Zhang, X. (1997). Prevalence of specific
47
48 language impairment in kindergarten children. *Journal of Speech, Language and Hearing*
49
50 *Research*, 40(6), 1245–1260.
51

52
53 Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children.
54
55 *Journal of Speech, Language, and Hearing Research*, 49(6), 1193.
56
57
58
59
60

Repetition of complex sentences in SLI

- 1
2
3 Tyack, D., & Ingram, D. (1977). Children's production and comprehension of questions. *Journal*
4
5
6 *of Child Language*, 4(2), 211–224.
7
8
9 Van der Lely, H. J. K., & Ullman, M. (2001). Past tense morphology in specifically language
10
11 impaired and normally developing children. *Language and Cognitive Processes*, 16(2-3),
12
13 177–217. <http://doi.org/10.1080/01690960042000076>
14
15
16
17 van der Lely, H. K. J. (2005). Domain-specific cognitive systems: insight from Grammatical-SLI.
18
19 *Trends in Cognitive Sciences*, 9(2), 53–59.
20
21
22 van der Lely, H. K. J., & Battel, J. (2003). Wh-movement in children with grammatical SLI: A test
23
24 of the RDDR hypothesis. *Language*, 79, 153 – 181.
25
26
27
28 van der Lely, H. K. J., Jones, M., & Marshall, C. R. (2011). Who did Buzz see someone?
29
30 Grammaticality judgement of wh-questions in typically developing children and children
31
32 with Grammatical-SLI. *Lingua*, 121(3), 408–422.
33
34
35
36 Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence - 3rd Edition*.
37
38 Pearson.
39
40
41
42 Wiig, E. H., Secord, W., & Semel, E. (1992). *Clinical Evaluation of Language Fundamentals -*
43
44 *Preschool*. San Antonio, TX: The Psychological Corporation.
45
46
47 Willis, C., & Gathercole, S. E. (2001). Phonological short-term memory contributions to sentence
48
49 processing in young children. *Memory*, 9(4), 349–363.
50
51
52
53 Wong, A. M.-Y., Leonard, L. B., Fletcher, P., & Stokes, S. F. (2004). Questions without
54
55 movement: A study of Cantonese-speaking children with and without specific language
56
57 impairment. *Journal of Speech, Language and Hearing Research*, 47(6), 1440.
58
59

Repetition of complex sentences in SLI

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Table 1

Group psychometric data – means and standard deviations (with z-scores in brackets)

	SLI (n= 17, 15 males)	AM (n = 17, 15 males)	LM (n= 21, 11 males)	Sig. difference on Tukey's test ($\alpha= .05$)
Age in months (years)	79.5 (6;7) 3.9	77.4 (6;5) 3.4	56 (4;8) 1.6	SLI > LM AM > LM
WPPSI non-verbal IQ	104.8 15.0	111.6 12.3	111.9 11.5	<i>no differences</i>
MLUw	6.6 1.2	8.0 0.8	6.7 1.2	AM > LM AM > SLI
CELF RS raw (z)	26.1 (-1.9) 10.5	44.7 (-0.2) 4.8	42.7 (0.9) 4.1	LM > SLI AM > SLI
CELF WS raw (z)	8.8 (-2.2) 2.9	16.2 (0.0) 1.7	12.8 (-0.1) 2.4	AM > LM AM > SLI LM > SLI
RAPT raw (z)	19.1 (-2.3) 5.0	26.7 (-0.1) 2.1	23.4 (0.7) 3.3	AM > LM AM > SLI LM > SLI
TROG blocks (z)	4.3 (-2.37) 2.1	10.1 (-0.45) 3.2	8.0 (0.65) 2.8	AM > SLI LM > SLI
BPVS raw (z)	58.2 (-0.44) 13.6	73.4 (0.49) 10.1	62.9 (1.07) 10.3	AM > LM AM > SLI
CNRep raw (z)	13.7 (-1.23) 7.5	24.8 (0.58) 5.0	17.8 (-0.35) 4.1	AM > SLI AM > LM

Table 2

Validity measures for dependent variables

Type of validity	Analysis	Dependent variable	
		Mean LDm per sentence	Number of repetition attempts containing word order errors
Construct validity	Effect of canonicity on dep. var. (LDm, word order errors) ^a	df = 17, t = -6.6 p < 0.001*** d = 0.52	df = 17, t = -3.0 p = 0.007** d = 1.4
	Association of dep. var. with elicitation task ^b	coeff. = -0.72 p = 0.002**	coeff. = 0.28 p = 0.302
Concurrent validity	Association with TROG	-0.57 p = 0.022*	0.44 p = 0.087
	Association with CNRep	-0.36 p = 0.168	0.42 p = 0.104

(a) Results of paired t-test with canonicity as the IV

(b) Results report Pearson's product moment correlations and respective significance values

(c) *** < 0.001, ** < 0.01, * < 0.05

Table 3

Errors by Group and Canonicity (mean and s.d.)

	SLI	Age-matched	Language-matched
Error rates in canonical sentences ^a	3.70 0.86	1.15 0.82	1.82 0.90
Error rates in non-canonical sentences ^a	4.46 0.84	1.55 1.06	2.36 0.78
Difference between conditions ^b	0.76 0.47	0.39 0.44	0.53 0.39
Difference expressed as a ratio ^c	1.22 0.16	1.39 0.47	1.51 0.59

(a) Mean number of errors per sentence per participant

(b) For each participant, the mean number of errors per canonical sentence was subtracted from the mean number of errors per non-canonical sentence, and then means and standard deviations of this measure were obtained.

(c) For each participant, the total number of errors in non-canonical sentences was divided by the total number of errors in canonical sentences, and then means and standard deviations of this measure were obtained

Table 4

Analyses of LDm using a zero-inflated negative binomial regression^a

Term	Coeff. (IRR) ^b	Lower / Upper 95% Conf. Int.	test statistic (z)	<i>p</i> (from z-test)	<i>p</i> (from LR test)
Group					< 0.001***
LM versus SLI	-0.86 (0.42)	-1.20 / -0.52	-4.90	< 0.001***	
AM versus SLI	-1.47 (0.23)	-1.84 / -1.11	-7.92	< 0.001***	
Canonicity	0.23 (1.26)	-0.05 / 0.50	1.60	0.11	0.06
Group x Canonicity					0.49
LM versus SLI	0.07 (1.07)	-0.05 / 0.19	0.88	0.38	
AM versus SLI	0.06 (1.07)	-0.08 / 0.21	1.10	0.27	
	Variance	St. Dev.			
Participant (int.)	0.27	0.52			
Canonicity (slope)	0.01	0.10			
Item (slope)	0.46	0.68			

(a) Random effects structure in R notation; (1+Canonicity|Participant) + (1|item)

(b) IRR = Incidence Rate Ratio

(c) *** < 0.001, ** < 0.01, * < 0.05

Table 5

Qualitative errors by Group

Percentage of repetition attempts exhibiting each error type (mean, s.d., range)

	SLI	AM	LM
Word order only errors ¹	1.9 (2.6)	0.4 (1.0)	0.9 (0.9)
Thematic role only errors ²	1.6 (1.5)	1.5 (2.2)	1.7 (1.8)
Combined word order and thematic role errors ³	6.1 (7.0)	3.6 (0.4)	4.5 (3.5)
Combined errors for canonical sentences	2.1 (2.9)	0.9 (1.9)	2.2 (3.0)
Combined errors for non-canonical sentences	10.2 (1.3)	6.4 (6.9)	6.7 (5.4)

1. e.g. *There's the cat that the dog chased* → *There's the dog that chased the cat*

2. e.g. *There's the cat that the dog chased* → *There's the dog that the cat chasde*

3. e.g. *There's the cat that the dog chased* → *There's the cat that chased the dog*

Table 6

Analysis of word order errors

	Fixed effects				
	Coeff.	SE	z	p (from z-test)	p (from LR test)
Group					0.144
LM vs SLI	0.18	0.54	0.33	0.741	
AM vs SLI	-0.91	0.65	-1.41	0.159	
Canonicity	1.77	0.42	4.24	< 0.001***	< 0.001***
Group x Canonicity					0.078
LM vs SLI	-0.43	0.48	-0.90	0.366	
AM vs SLI	0.44	0.58	0.751	0.453	
Random effects ^a					
	Mean	Variance			
Participant	1.36	1.17			
Item	1.91	1,38			

(a) A fully-specified random effects model was fitted. This did not converge. The parameter for the correlation between Participant (intercept) and Canonicity (slope) was removed (Baayen et al. 2008), but the model still did not converge. Finally, a model removing Canonicity as a random slope, but retaining random intercepts for Participant and Item was fitted; (1|participant) + (1|item). This final model converged and is shown above. In addition, a model with (1+canonicity|participant) was also run, as proposed by Barr et al. (2013) who argue that it is important to run random slopes models in cases of non-convergence. This model yielded identical results in terms of significance.

(b) *** < 0.001, ** < 0.01, * < 0.05

Table 7

Analysis of SLI subgroup

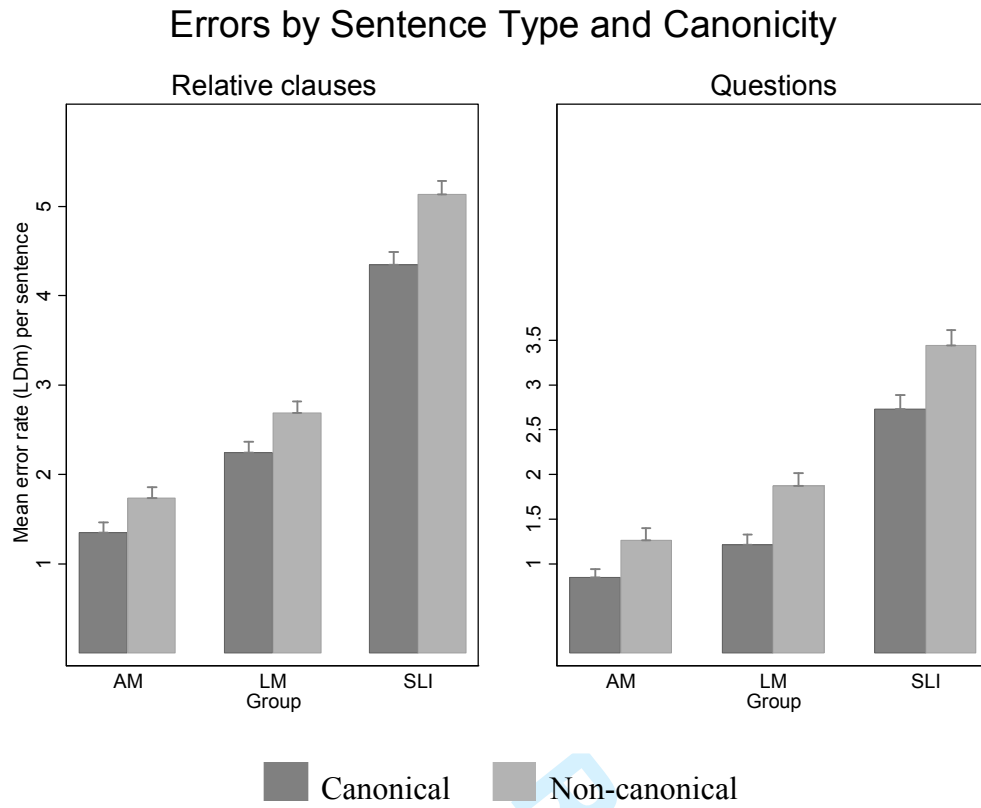
Language measure	Scores for participants 1, 3 and 11 respectively	Comparison with group	Mean for SLI group (and s.d.)
SR Ratio ^a	1.27		1.23
	1.55	(> +1 s.d.)	0.16
	1.51	(> +1 s.d.)	
SR Difference ^b	1.02		0.76
	1.58	(> +1 s.d.)	0.47
	0.88		
Errors per sentence (LDm)	3.7		4.1
	2.9	(< -1 s.d.)	0.82
	4.5		
Score on elicitation task	1	(> +1 s.d.)	0.63
	0.61		0.32
	0.60		
Score on RAPT	22		17.6
	25	(> +1 s.d.)	6.7
	18		

(a) Ratio = number of errors in non-canonical sentences per error in canonical sentences

(b) Difference = mean error rate for non-canonical sentence minus mean error rate per canonical sentence

First three measures are negatively scored with high scores denoting poor performance, while second two measures are positively scored with high scores denoting good performance. Bold values show performance substantially worse than group

Figure 1



Error bars show standard error of the mean

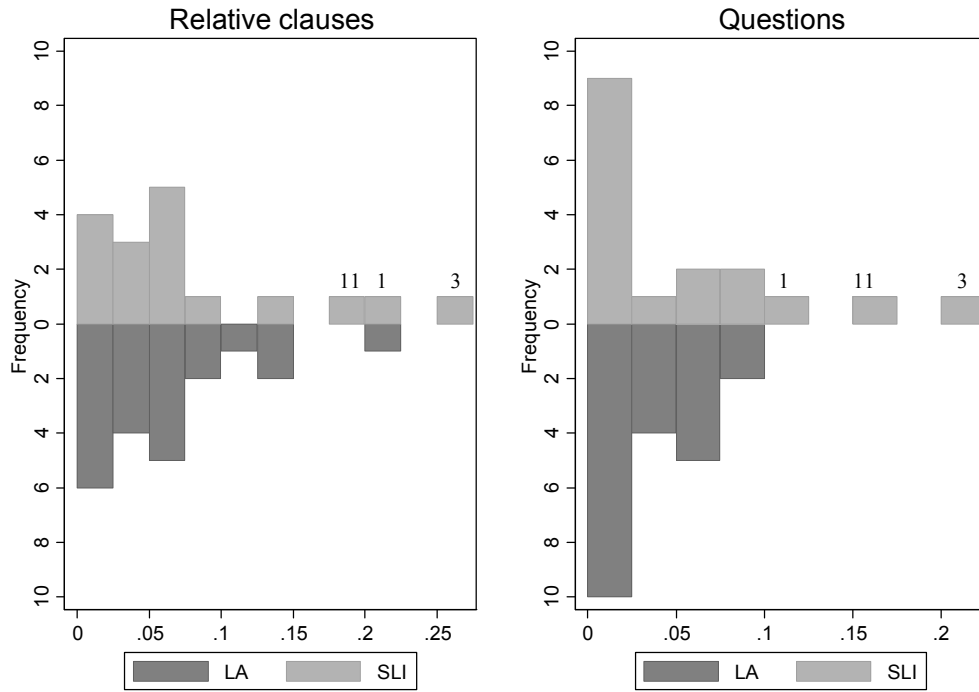
Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 2

Histograms of word order errors by group and structure

Proportion of repetition attempts changing word order, but maintaining serial order of NPs



Numbers show ID numbers for outlying children with SLI