Published in final edited form as:
J Psycholinguist Res. 2010 October ; 39(5): 443–456. doi:10.1007/s10936-010-9151-1.

Relative Clause Gap-Filling in Children with Specific Language Impairment

Arild Hestvik,
Department of Linguistics and Cognitive Science, University of Delaware, 42 East Delaware Avenue, Newark, DE 19716, USA

Ph.D Program in Speech-Language-Hearing Sciences, Graduate Center, City University of New York, New York, NY, USA

Richard G. Schwartz, and
Ph.D Program in Speech-Language-Hearing Sciences, Graduate Center, City University of New York, New York, NY, USA

Lydia Tornyova
Department of Linguistics, Graduate Center, City University of New York, New York, NY, USA

Abstract

Children with Specific Language Impairment (SLI) have been observed to have production and perception difficulties with sentences containing long-distance dependencies, but it is unclear whether this is due to impairment in grammatical knowledge or in processing mechanisms. The current study addressed this issue by examining automatic on-line gap-filling in relative clauses, as well as off-line comprehension of the same stimulus sentences. As predicted by both knowledge impairment and processing impairment models, SLI children showed lack of immediate gap-filling after the relative clause verb, in comparison to a control group of typically developing children. However, on the off-line measure of comprehension of the same stimuli sentences, SLI children and TD children did not differ qualitatively. This finding is incompatible with knowledge impairment. We interpret the results to show that SLI children have impaired processing mechanisms (such as temporally delayed gap-filling) but are not impaired in their grammatical knowledge.

Keywords

Gap-filling; Antecedent reactivation; SLI; Language impairment

Introduction

Children with Specific Language Impairment (SLI) have difficulties comprehending and producing sentences that involve long-distance dependencies, such as Wh-questions (Deevy and Leonard 2004; Van der Lely and Battell 2003) and relative clauses (Friedmann and Novogrodsky 2004; Håkansson and Hansson 2000; Novogrodsky and Friedmann 2006; Schuele and Nicholls 2000; Stavrakaki 2001). According to (Van der Lely 2004, 2005; Van der Lely and Battell 2003; van der Lely et al. 1998), the source of this difficulty is an impairment in SLI children's innate predispositions towards grammatical induction—an

arild@udel.edu .

impairment in Universal Grammar (UG). This view holds that SLI children construct grammars which fail to represent long-distance dependencies (or where the dependency is "optionally" represented). Thus, in a relative clause like (1), there is no guarantee that the relationship between the filler *the zebra* and the gap after the verb will be established:

(1) *The zebra* that the hippo on the hill had kissed *t* on the nose ran far away

If SLI children fail to fill the gap in sentences like (1), they should not exhibit an increase in activation of the filler in this position, and they should have severe comprehension problems with such sentences (in fact, complete failure of comprehension). On this view, SLI children have a different grammar than their typically developing peers, but their processing system is unimpaired and implements whatever grammar it is that they have induced. Because the difference between SLI children and their typically developing peers resides in the grammar, not in their processing system, SLI children's gap-filling should be equally impaired under varying processing demands.

An alternative view is that SLI children do develop the same basic grammatical knowledge as TD children (although at a slower pace), but that their problems are related to a grammar-independent deficit in real-time language comprehension mechanisms (Leonard 1998). This predicts that performance should vary with processing demands that interact with the impairment. Factors such as sentence length (Deevy and Leonard 2004) and verbal working memory demands (Montgomery 2000, 2003) have been shown to modulate SLI children's performance. These findings are inconsistent with knowledge impairment, because processing demands by themselves are insulated from grammatical knowledge under the assumption of a competence-performance distinction.

Additionally, knowledge impairment predicts that comprehension of sentences requiring gap-filling should be severely impaired. A processing impairment does not necessarily entail this. For example, the impairment could be a temporally delayed gap-filling process. This would predict a lack of gap-filling in the immediate temporal vicinity of the relevant verb, but does not rule out that SLI children can perform "delayed" gap-filling, or process the dependency via alternative or additional strategies.

The current study aimed to differentiate between knowledge impairment and processing impairment by examining the relationship between on-line and off-line processing of relative clauses. We examined (A) whether SLI and TD children differed in their real-time construction of filler-gap dependencies in relative clauses (as predicted by both the knowledge and processing account of SLI), and (B) whether the two groups also differed in their comprehension of the stimuli sentences, as measured by off-line comprehension questions. Based on a previous study of SLI gap-filling in Wh-questions (Marinis and Van der Lely 2007), we expected to observe lack of gap-filling in relative clauses as well. This by itself would not distinguish between knowledge impairment. Additionally, a finding of reduced off-line comprehension would also not help in distinguishing between the two models. However, relatively good comprehension in the absence of immediate gap-filling would allow for the rejection of knowledge impairment and lend support to a processing impairment model.

The Current Study: Cross-Modal Picture Naming

Examining gap-filling in children and especially in language impaired child populations presents methodological challenges. Gap-filling has been measured in adults with the cross-modal priming technique, by which subjects make lexical decisions to visual word probes appearing at gap positions and control positions during auditory sentence presentation (Love and Swinney 1996; Nicol et al. 1994b; Nicol and Swinney 1989; Swinney et al. 1989). If a gap has been filled, the lexical information associated with the filler will prime probes presented at this temporal position. Selective priming at the gap position compared to no priming at a pre-gap control condition is taken as evidence of gap-filling has taken place.

This method has also been used with children, by replacing lexical decision tasks with categorical decision tasks for picture probes (Love 2007; Love and Swinney 1997; McKee et al. 1993; Roberts et al. 2007). Recently, this paradigm has been used to demonstrate lack of gap-filling for Wh-questions in SLI children (Marinis and Van der Lely 2007). The goal of the current study was to measure gap-filling in relative clauses in SLI children. However, we reasoned that it was necessary to construct a task that had the minimum of cognitive resource requirements, in order to prevent a task bias against the SLI children. We therefore adapted the design and stimuli of Love (2007), but instead of requiring the children to make binary categorical decisions for each picture, we asked them to simply *name* the picture as fast as they could. The rationale was that picture naming is easy and effortless for children, and require less processing resources that binary categorical decision tasks. We also reasoned that picture naming might yield greater magnitude of priming, because it involves a direct, lexical priming relationship, in comparison to the indirect, semantic priming in categorical and lexical decision tasks. The rationale was that if the effect was small and the SLI data were noisy, a larger priming effect would help to overcome such obstacles.

Thus, in our paradigm, children hear relative clauses like (1), and a picture of either the relativized noun or a control picture is visually presented at the gap position, or at a pre-gap control position, in a 2 (position) × 2 (probe type) design. For example, children heard sentences like (1) with a picture of a zebra temporally displayed at the offset of the verb (position marked by subscript 2).

(1) *The zebra* that the hippo on the hill$_1$ had kissed$_2$ on the nose ran far away

If the gap is immediately filled by "zebra" at this point in time, naming of a picture of a zebra should be facilitated in comparison to naming of a control picture. To control for the possibility that the priming is merely a continuation effect, or a repetition priming effect from the initial relativized noun, the same two picture probes was also presented in the pre-gap control position (marked by subscript 1). In the pre-gap control position, no priming was expected.

In order to draw inferences from this study, we first tested in a separate experiment the assumption that SLI and TD children all have basic word-picture priming, independently of gap-filling and reactivation.

Experiment 1

Experiment 1 constituted a pre-test of the name-picture pairings, to ensure that both SLI children and TD children satisfied a pre-condition for interpreting the results of the main experiment, namely, basic word-picture priming independently of syntactic structure and sentence embedding. At the same time, Experiment 1 served to familiarize the participants with the pictures and their names.

Method

Subjects—A total of 60 children were initially enrolled in the study. 1 TD subject was excluded because of experimenter error; 2 SLI and 10 TD were excluded because one or more parts of the protocol was incomplete, 2 SLI and 2 TD children were excluded because of unclear clinical status, and 3 TD subjects were excluded because of too high IQ scores (over 140). The high rate of attrition was due to long duration of the study (4 experimental session and additional visits for screening tests). This left 20 SLI children and 20 TD children with complete data. None of the children had hearing impairment or a history of neurological problems. All SLI children had been referred to services by a school speech therapist.

The two groups of children were matched on age, IQ and verbal memory span, as confirmed by statistical tests of no significant difference between the groups on these measures. (All the reported standardized tests have a mean of 100 and a standard deviation of 15.)

The mean age of the SLI children was 9.7 (SD = 1.5, range 8–13 years; 6 girls and 14 boys), and the mean age of the TD children was 9.4 (SD = 1.3, range 8–12 years; 13 girls and 7 boys). The mean IQ for the SLI group as measured by the Test of Non-Verbal Intelligence, TONI-3 (Brown et al. 1997) was 102.7 (SD = 13.2); all SLI children had IQ within normal range. The mean IQ for the TD group was 107.7 (SD = 14.1). A one-way ANOVA showed the difference to be non-significant ($F(1, 38)$ 1.= 4, $p = .25$). Verbal memory span was measured by the Competing Language Task (Gaulin and Campbell 1994), a child adaptation of the Listening Span Test for adults (Daneman and Carpenter 1980). The mean verbal memory span score for the SLI group was 2.48 (SD = 0.57), and the mean span for the TD group was 2.45 (SD = 0.87), a non-significant difference by one-way ANOVA ($F(1, 38) = .011$, $p = .92$). Hence, no observed experimental difference between the groups could be attributed to general intelligence, age, or verbal memory span differences.

On the Peabody Picture Vocabulary Test, PPTV-4 (Dunn and Dunn 2007), the mean score for the SLI group was 91.1 (SD = 8.5) and the mean score for the TD group was 106 (SD = 12.7), a significant difference ($F(1, 38)=19.4$, $p < .001$). On the Comprehensive Evaluation of Language Fundamentals test, CELF-4 (Semel et al. 2004), the expressive language score for the SLI group was 79.9 (SD = 10.6), and the composite language score was 81.5 (SD = 11.7). The corresponding scores for the TD group were 117 (SD = 12.5) on the expressive test and 115 (SD = 11.6) on the composite score. Thus, the SLI group scored below 1.3 SDs below the mean on a set of language screening tests, a standard criterion for SLI in the literature.

Materials and Design—Sixteen animal names (“zebra”, “hippo”, “camel”, etc.) were recorded by a female speaker for auditory presentation, and matched to pairs of pictures drawn from (Rossion and Pourtois 2001), where one picture depicted the named animal (the “primed” picture) and the control picture depicted another animal (but from the same set of names). For example, the noun “zebra” would be matched to the two pictures in Fig. 1.

The primed picture was in turn used as a control picture for a sentence where the control picture for zebra was the primed picture (i.e., a sentence with “camel” as filler). This “matched probe” design controls for item effects and obviates the need for conducting an ANOVA with item as a random factor. Each picture name was recorded as a single word, and matched to pairs of pictures: the primed picture and a designated control picture (as in Fig. 1).

Procedure—The procedure in Experiment 1 was as follows: Each of the sixteen animal names were presented via headphones in isolation, followed by a 500 ms inter-stimulus interval, and then followed by the primed picture or the control picture (on separate trials). Trials were presented in the same pseudo-randomized order for each subject. An additional 24 filler trials, consisting of other animal names but matched with inanimate pictures, were interspersed with the experimental trials. The children were told to listen to the presented word and then name the picture presented on a computer screen as fast as they could. Voice response latency was measured by voice key (PST Serial Response box), and naming accuracy was manually scored. The picture was presented for a maximum of 2000 ms, and was automatically removed from the screen if naming occurred before time-out. Each trial was presented twice by repeating the experiment in a second session about a week later, for a total of 64 experimental trials.

Results—Mean naming accuracy was computed for each subject by condition, and was used as the dependent measure in a 2 (probe type) × 2 (group) repeated measures ANOVA. A main effect of probe type was observed ($F(1, 38) = 20.3$, $p < .0001$), such that the mean naming accuracy was higher for primed pictures (M = 89%, SD = 7%) than for control pictures (M = 82%, SD = 11%). A main effect of group was also observed ($F(1, 38) = 7.5$, $p < .01$), such that the mean accuracy was higher for TD children (M = 89%, SD = 10%) than for SLI children (M = 83%, SD = 7%), a 6% difference. There was no interaction between group and probe type.

In addition, mean naming accuracy was computed for each picture. This revealed two picture items that had low overall naming accuracy (less than 75% mean accuracy in each group), which coincided with experimenter observations. In particular, the picture for "bee" was frequently named "fly" (71% accuracy across priming conditions for the SLI group, 64% accuracy for the TD group); and the picture corresponding to the prime word "kitten" was mostly named "cat" (only named "kitten" 25% of the time by the SLI children and 71% of the time by the TD children). Finally, the control picture for "horse", which was a picture of a donkey, was sometimes named "horse" or lead to hesitations. This is due to the fact that hearing "horse" primes for a horse picture, and a donkey obviously looks a bit like a horse. However, removing the items for "bee" and "cat" did not change the statistical results.

For the reaction time data, trials that had been manually scored as false starts, naming errors and voice trigger errors (14% of the total number of experimental trials; the breakdown was 17% for SLI and 11% for TD) were removed from analysis. An additional 16 trials (out of 2193) that were faster than two standard deviations from the mean were treated as errors and also removed. The mean naming latency per subject by condition was then computed and used as the dependent measure in a 2 (probe type) × 2 (group) repeated measures ANOVA. This resulted in a main effect of probe type ($F(1, 38) = 136$, $p < .0001$), such that primed pictures were named faster (M = 658 ms, SD = 139) than control pictures (M = 812 ms, SD = 120). There were no other main effects or interactions between group and priming. The mean naming latency for SLI children was numerically higher than the TD children by 57 ms, but this was not a significant difference ($F(1, 38) = 2.2$, $p = .15$). Removal of the two items with low naming accuracy did not affect the results.

Discussion—Experiment 1 showed that both TD and SLI children exhibited a strong priming effect when the critical animal names and the associated picture probes were presented in isolation. Thus, a pre-condition for interpreting the results in Experiment 2 is met, because any difference between the two groups cannot be explained as SLI children not having a basic word-picture priming effect. SLI children had significantly lower naming accuracy than TD children, but the difference was small (6%). Most importantly, there was no difference between the two groups with respect to the basic priming effect.

Experiment 2

Method

Subjects—The same subjects who participated in Experiment 1 also participated in Experiment 2 about a week later (each visit for the study was kept to about a week apart).

Materials and Design—For experiment 2, each animal name was embedded as relativized object nouns in relative clauses, and the relative clause itself modified the subject of a sentence. Eight of the sixteen sentences were first assigned to the "gap" condition, i.e.

presented with picture probes at the temporal position of the gap. The other eight sentences were assigned to the pre-gap control condition, presented with picture probes at the offset of a prepositional phrase modifier after the relative clause subject (“Script A”). The “matched probe” and probe position design structure is illustrated in (2):

(2) ScriptA, pre-gap control condition:

- **a** The zebra that the hippo on the hill$^{zebra/camel}$ had kissed on the nose ran far away. The camel that the rhino in the mud$^{zebra/camel}$ had watched from a distance peeked through the leaves.

Script A, gap condition:

- **b** The bear that the gorilla in the mist had scared$^{bear/tiger}$ by accident went swimming in the pool. The tiger that the lion at the rock had followed$^{bear/tiger}$ in the woods hid behind a tree.

In addition, a second set of stimuli (“Script B”) was constructed by exchanging the relative clause verb and main clause verb phrase in each sentence with that of another sentence, thus doubling the set of sentences. In this version of the sentences, the word that primed a picture in the pre-gap control position in the first set now also primed a picture in the gap position, and vice versa:

(3) Script B, gap condition:

- **a** The zebra that the hippo on the hill had watched$^{zebra/camel}$ from a distance peeked through the leaves. The camel that the rhino in the mud had kissed$^{zebra/camel}$ on the nose ran far away.

Script B, pre-gap control condition:

- **b** The bear that the gorilla in the mist$^{bear/tiger}$ had followed in the woods hid behind a tree. The tiger that the lion at the rock$^{bear/tiger}$ had scared by accident went swimming in the pool.

Thus, each of the 16 relativized nouns were matched once to each of the 2 × 2 cells of the probe type × probe position design, and in a unique sentence for each probe position. For example, the relativized noun “zebra” would occur in a sentence with a control probe in the gap position, a primed probe in the gap position, as well as control probe in the control position and a primed probe in the control position.

In addition, 37 additional animal names and sentences were constructed to be used as filler trials. These sentences had various syntactic structures. Each filler sentence was matched to randomly sampled pictures of inanimate objects, and picture positions were varied across different temporal positions, so that a picture would seem to the subject to appear at any time during a sentence, throughout the experiment. The inanimate picture for the filler trials were randomly selected from a set of filler pictures each time a filler sentence was presented.

Both the probe type and probe position factors were within-subject, with group (SLI vs. typical developing) as the between-subjects factor, thus constituting a mixed factorial repeated measures design.

Procedure—Each of the 16 experimental sentences was distributed across multiple sessions so that only one version of each sentence in the 2 × 2 design was presented on a given session, but each relativized noun was presented twice in each session (in a different sentence). i.e., on one session, half the sentences had probes in the control position and half in the gap position, crossed with probe type so that half the sentences were presented with control probes and half with primed probes. For example, a sentence with relativized "zebra" would be presented once in the first session with a primed picture in the pre-gap control position, and once in another script sentence in the gap position. On the second session, the control probes would be presented for each position. All trials during one session were presented in pseudo-randomized order.

Note that this also means that subjects saw the same picture of a zebra twice in the same session, once in the control position and once in the gap position. This would give a naming advantage for the picture on its second exposure. However, order of presentation was pseudo-randomized so that half the pictures were first presented in the control position and half was first presented in the gap position. Furthermore, half the pictures functioned as control pictures in both positions and the other half functioned as primed pictures in both positions, within a session. In the next session, these roles were switched. Thus, any advantage on naming latency of seeing the primed probe a second time would be counter-balanced by an equal advantage on naming latency on seeing the control probe a second time.

In order to maximize the probability of a measurement in each cell of the 2 × 2 design for each test sentence, and avoid missing data in the repeated measures design at the item level per subject, each sentence was presented twice in each cell throughout the course of the study. This therefore required four sessions. Script order, probe type order (i.e. related or control first for a given item) and probe position for each sentence was fully counterbalanced across all subjects and sessions. In addition, 48 filler sentence trials were presented during each session (requiring11 filler sentences to be repeated in each session). Inanimate pictures were randomly sampled and presented for each filler trial. Thus, at the end of the study, every subject had heard every sentence twice in each of the four cells, totaling 128 experimental trials and 192 filler trials, over four experimental sessions (conducted roughly one week apart at a minimum).

The sentences were presented through headphones at normal speaking rate by a female voice. Children were instructed to pay attention to the meaning of the sentence, and name the picture when it appeared on the computer screen as fast as they could, while maintaining attention to sentence meaning. The picture remained on the screen for a maximum of 3000 ms, but was removed before this time-out if a naming response occurred. (The response window in Experiment 2 was chosen to be 1000 ms longer than in the single word-picture presentation in Experiment 1 because we expected that the greater demands of processing entire sentences and answering comprehension questions would lead to overall longer response times.)

Each trial was followed by a comprehension question. The comprehension questions were of three types: a question about the matrix subject, about the relative clause subject, or about the relative clause object. For example, for a sentence like (1), the following three questions could be asked:

(4) **a.** Who ran far away?

b. Who had kissed the zebra?

c. Who had the hippo kissed?

Even though the answer to both (4a) and (4c) is "zebra", (4a) is labeled a matrix subject question because it uses the main verb, whereas (4c) uses the relative clause verb. Picture naming accuracy and comprehension question accuracy was scored manually, and naming latencies were recorded by voice key.

Results—Each subject's mean picture naming accuracy was computed for probe type and probe position, and used as dependent measures in a 2 (probe type) × 2 (position) × 2 (group) mixed factorial repeated measures ANOVA. This revealed a main effect of group ($F(1, 38) = 5.0$, $p < .05$), such that mean naming accuracy for TD children (89%, SD = 10%) was higher than for SLI children (84%, SD = 7%), and a main effect of probe type ($F(1, 38) = 31.1$, $p < .0001$), such that subjects had higher naming accuracy for primed pictures (90%, SD = 6%) than for control pictures (83%, SD = 8%). There were no further main effects or interactions.

As in Experiment 1, trials with naming errors, false starts and voice key trigger errors, as well as response times faster than 2 standard deviations from the mean were removed; this constituted 13% of all trials (690 out 5120). Secondly, we removed all remaining trials for which the subject gave the wrong answer to the comprehension question. This is critical as it is very easy to disregard the sentence and just focus on naming the picture (which children evidently did a good portion of the time). On a trial where the sentence material is "tuned out", picture naming latency is not likely to reveal much about the processing of the auditory sentence. Out of the total number of trials, subjects answered incorrectly 25% of the time, and out of the trials with correct picture naming, subjects answered incorrectly 30% of the time. This bears witness to the fact that the experiment was very demanding for the children, and that performing correctly on the divided attention task posed even a greater challenge. Nevertheless, due to the high trial count this still leaves a sufficient number of data points for statistical analysis. After removing trials with wrong comprehension answers in this way, the SLI children had on average 19 (SD = 3.4) completely error-free trials in each of the four conditions, and the TD children had on average 20 (SD = 3.8) error-free trials per condition. A power analysis of a test of priming in one probe position within a subject was conducted. With a sample size of 20 items for a paired t-test with effect size 0.5, there is a 70% chance of detecting an difference.

We next examined picture naming accuracy by item. The average naming accuracy for primed pictures was 94% (SD 9%) and the average naming accuracy for pictures functioning as control was 89% (SD 10%); however, items using the picture intended to be named as "kitten" were named correctly only 63% and 54% respectively. For this reason, trials with the two stimuli sentences using this picture item as primed was removed from analysis, and due to the matched probe design, so were two sentences using this as a control picture (matched to prime word "pig"). The remaining trials which had both correct picture naming and correct comprehension question answers were then used to compute a mean accuracy per subject for each probe type and probe position.

We finally calculated the mean comprehension accuracy score for each subject, i.e. proportion correctly answered comprehension question for all trials presented, irrespective of whether the picture was correctly named or not. If a subject's accuracy on the comprehension questions were consistent with chance (i.e. within the 95% confidence interval of 50%), they were classified as guessing, i.e. not attending sufficiently to the meaning of the sentence, and excluded from analysis. In other words, whereas we first removed trials based in incorrect comprehension questions, we now removed subjects based on chance performance on the comprehension questions. Exactly 6 SLI subjects and 6 TD subjects were excluded on the basis of this criterion.

For the remaining 14 subjects in each group, mean reaction times in each of the four 2 × 2 cells were used as the dependent measure in a mixed factorial repeated measures design. These means are illustrated in Fig. 2:

The omnibus ANOVA resulted in a main effect of probe position ($F(1, 26) = 16.3$, $p < .001$) such that pictures were named faster at the gap position than the earlier control position. A main effect of probe type was also observed ($F(1, 26) = 61.1$, $p < .00001$), such that primed pictures were named faster than control pictures. No other main effects or interactions were observed. This is not to say that the there were no differences between the two groups, but rather that differences were too subtle to be detected by an omnibus 3-way ANOVA. Because the experiment was designed to examine differences between priming at the two positions within each group, we next conducted Bonferroni-protected t-tests, and interpreted those comparisons that were motivated by the design of the experiment. An additional motivation for using Bonferroni-correction is that this experimental paradigm is novel and we therefore cannot base predictions on past experimental results. The means, standard deviations and two-tailed probabilities for each comparison are shown in Table 1.

For the TD children, these comparisons showed that the difference between control and primed picture probes were not significant in the control position, but were significantly different in the gap position. For the SLI group, the opposite pattern was observed: The difference between primed and control pictures were significant in the pre-gap control position, but not at the gap position. In addition, a 2 × 2 repeated measures ANOVA for the TD group with probe and position as crossed factors yielded main effects of probe ($F(1, 13) = 21.8$, $p < .001$) and position ($F(1, 13) = 7.1$, $p < .05$, but the interaction term was not significant in this analysis ($F(1, 13) = 0.9$, $p = 0.34$.

We next performed a detailed analysis of the off-line performance on the comprehension questions. In this analysis, we examined whether subjects' accuracy on comprehension questions depended on the type of comprehension question asked, and whether it depended on the experimental condition preceding the comprehension question. Thus, we now analyze all trials and use comprehension as a dependent measure. As outlined above, each trial was followed by either a matrix subject question, or a question about the subject of the relative clause (using the relative clause verb), or a question about the relativized object—the filler. We computed the mean accuracy per subject for each question type on only trials with correct picture naming (again excluding the trials associated with the low-accuracy "kitten" picture), and conducted a 3 (question type) × 2 (group) repeated measures factorial ANOVA. This revealed a main effect of question type ($F(2, 52) = 28.3$, $p < .0001$), but no other main effect or interactions. Mean comprehension question accuracy for the 14 TD children was 78% (SD 11%), and 76% (SD = 12%) for the SLI children; these means are summarized in Fig. 3.

Scheffé posthoc tests showed that all three types of comprehension question differed from each other (matrix subject questions: 77%; relative clause object questions: 88%; relative clause subject questions: 66%).

Because half of the comprehension questions about matrix subject and relative clause object had an answer that coincided with a primed picture, we examined whether groups differed in terms of whether they were "helped" by priming at answering comprehension questions. We categorized the questions by whether the answer was primed or not, and conducted a 2 (primed answer) × 2 (group) ANOVA. This revealed a main effect on comprehension question accuracy by primed picture ($F(1, 26) = 37.4$, $p < .0001$), but no other effects. Finally, the question type that never had a primed answer (relative clause subject questions) was also analyzed separately with a one-way ANOVA. This question type showed no difference between the two groups ($F(1, 26) = .07$).

Discussion—The pair-wise comparisons of the reaction time means in Experiment 2 show that the TD children exhibited significant priming in the gap position, but not so in the pre-gap control position. SLI children on the other hand, showed the opposite pattern: They exhibited significant priming in the pre-gap control position, but no significant priming at the gap. We interpret this to show that TD children but not SLI children have immediate gap-filling after the verb. Note that the effect in the TD group was not so robust as to yield a significant interaction effect in a 2 × 2 ANOVA, which has been argued by (Nicol et al. 1994a) to be the strongest evidence for reactivation. However, our a priori predictions of significant priming at the gap in comparison to the pre-gap position were met and demonstrate that TD children but not SLI children exhibit the predicted change in activation of the filler at the gap.

In addition, there were also main effects distinguishing between the two groups: SLI children were significantly slower at naming pictures, and significantly less accurate. Nevertheless, both groups showed robust main effects of priming in both experiments. Turning to the off-line measure, there was no difference between groups on a measure of accuracy of comprehension: TD and SLI children had the same comprehension scores and showed the same pattern of accuracy depending on the grammatical structure of the comprehension questions.

Conclusion

The reaction times findings indicate that when processing a relative clause, TD children are similar to adults in that they immediately fill the gap at the pre-verbal position, whereas SLI children fail to show immediate gap-filling. This finding represents a methodological advance, as it shows that cross-modal picture naming replicates previous studies that used cross-modal picture categorizations tasks for studying gap-filling in relative clauses and Wh-questions (Love 2007; Roberts et al. 2007). In addition, the finding of a lack of immediate reactivation after the verb in SLI children converges with a similar recent finding for Wh-questions in SLI children (Marinis and Van der Lely 2007). Thus, there is now strong evidence from different types of long-distance dependencies that SLI children fail to show immediate reactivation effects in gap-filling constructions.

We have no account to offer for why SLI children showed priming at the pre-gap control position. It is possible that the picture naming task contains an additional component of repetition priming that would boost priming at the pre-gap position in general, and that this shows up in SLI children more than in TD children.

As discussed in the Introduction, lack of reactivation is consistent with both a knowledge impairment account and a processing impairment account of SLI. However, the finding that there was no difference between the two groups on the off-line comprehension measure, despite a difference in reactivation at the gap, is inconsistent with knowledge impairment. If SLI children fail to construct a gap after the verb, they should exhibit a severe comprehension deficit, or at least, significantly lower comprehension score of the relative clauses as measured by comprehension questions. The fact that SLI children are virtually identical to the TD children on the off-line comprehension measure suggest that they are able to assign a correct meaning to the stimulus sentence, which means that they must be able to establish the relationship between the filler and the verb. This is consistent with theories that propose that the core deficit in SLI is impairment in language processing functions, rather than in grammatical knowledge (Leonard 1998). To speculate, one possibility is that the lack of immediate reactivation in SLI children is due to an attenuation of the Active Filler Strategy (Clifton and Frazier 1986, 1989; Frazier and Flores d'Arcais 1989; Stowe 1986). In the domain of gapfilling, the Active Filler strategy entails that the parser makes fast and immediate predictions about where the gap for a filler should be located. The current findings would be consistent with a model where the SLI parser does

NIH-PA Author Manuscript

NIH-PA Author Manuscript

not predict gap positions with the same speed as a typical parser. If this were the case, SLI children do establish the filler-gap dependency, but do so at a slower rate. In particular, the temporal juncture at which the cross-modal priming method measures reactivation is “too soon” for the SLI children, which accounts for a lack of reactivation effect at this position. However, if they eventually do establish the relationship, that would account for the similar off-line comprehension scores observed in the current study. This raises the question of whether reactivation could be observed in SLI at later temporal positions, which we leave as an open question for future research.

Acknowledgments

This research was supported by NIH grant 5R03DC006175 from the National Institute on Deafness and Other Communicative Disorders (NIDCD) to the first author, and by NIH grant 5R01DC003885-03 from NIDCD to Richard G. Schwartz.

References

Brown, L.; Sherbenou, RJ.; Johnsen, SK. Test of nonverbal intelligence. 3rd ed.. Pro-Ed; Austin, TX: 1997.

Clifton C Jr. Frazier L. The use of syntactic information in filling gaps. Journal of Psycholinguistic Research 1986;15(3):209–224. [PubMed: 3723441]

Clifton, C., Jr.; Frazier, L. Comprehending sentences with long-distance dependencies.. In: Carlson, G.; Tanenhaus, MK., editors. Linguistic structure in language processing. Kluwer Academic Publishers; Dordrecht: 1989. p. 273-317.

Daneman M, Carpenter P. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior 1980;19:450–466.

Deevy P, Leonard LB. The comprehension of wh-questions in children with specific language impairment. Journal of Speech, Language, and Hearing Research 2004;47(4):802–815.

Dunn, LM.; Dunn, DM. Peabody picture vocabulary test. 4th ed.. Pearson Inc.; Bloomington, MN: 2007.

Frazier L, Flores d’Arcais GB. Filler driven parsing: a study of gap filling in dutch. Journal of Memory and Language 1989;28(3):331.

Friedmann N, Novogrodsky R. The acquisition of relative clause comprehension in Hebrew: a study of SLI and normal development. Journal of Child Language 2004;31:661–681. [PubMed: 15612394]

Gaulin CA, Campbell TF. Procedure for assessing verbal working memory in normal school-age children: some preliminary data. Perceptual and Motor Skills 1994;79(1Pt1):55–64. [PubMed: 7991333]

Håkansson G, Hansson K. Comprehension and production of relative clauses: a comparison between Swedish impaired and unimpaired children. Journal of Child Language 2000;27(02):313–333. [PubMed: 10967890]

Leonard, LB. Children with specific language impairment. MIT Press; Cambridge, Massachusetts: 1998.

Love T. The processing of non-canonically ordered constituents in long distance dependencies by pre-school children: a real-time investigation. Journal of Psycholinguistic Research 2007;36(3):191–206. [PubMed: 17186385]

Love T, Swinney D. Coreference processing and levels of analysis in object-relative constructions: demonstration of antecedent reactivation with the cross-modal priming paradigm. Journal of Psycholinguistic Research 1996;25(1):5–24. [PubMed: 8789365]

Love, T.; Swinney, D. Real time processing of object relative constructions by pre-school children [Poster].. 10th Annual CUNY conference on human language processing.; Santa Monica, CA. 1997.

Marinis T, Van der Lely H. On-line processing of wh-questions in children with G-SLI and typically developing children. International Journal of Language & Communication Disorders 2007;42(5): 557–582. [PubMed: 17729146]

McKee C, Nicol J, McDaniel D. Children's application of binding during sentence processing. Language and Cognitive Processes 1993;8:265–290.

Montgomery JW. Verbal working memory and sentence comprehension in children with specific language impairment. Journal of Speech, Language, and Hearing Research 2000;43(2):293–308.

Montgomery JW. Working memory and comprehension in children with specific language impairment: what we know so far. Journal of Communication Disorders 2003;36(3):221–231. [PubMed: 12742669]

Nicol J, Fodor J, Swinney D. Using cross-modal lexical decision tasks to investigate sentence processing. Journal of Experimental Psychology: Learning, Memory & Cognition 1994a;20(5):1–10.

Nicol J, Fodor JD, Swinney D. Using cross-modal lexical decision tasks to investigate sentence processing. Journal of Experimental Psychology: Learning, Memory and Cognition 1994b;20(5): 1229–1238.

Nicol J, Swinney D. The role of structure in coreference assignment during sentence comprehension. Journal of Psycholinguistic Research 1989;18(1):5–19. [PubMed: 2647962]

Novogrodsky R, Friedmann N. The production of relative clauses in syntactic SLI: a window to the nature of the impairment. Advances in Speech-Language Pathology 2006;8:364–375.

Roberts L, Marinis T, Felser C, Clahsen H. Antecedent priming at trace positions in children's sentence processing. Journal of Psycholinguistic Research 2007;36(2):175–188. [PubMed: 17160507]

Rossion B, Pourtois G. Revisiting Snodgrass and Vanderwart's object database: color and texture improve object recognition. Journal of Vision 2001;1(3):413a.

Schuele CM, Nicholls L. Relative clauses: evidence of continued linguistic vulnerability in children with specific language impairment. Clinical Linguistics & Phonetics 2000;14(8):563–585.

Semel, E.; Wiig, EH.; Secord, WA. Clinical evaluation of language fundamentals. 4th ed.. The Psychological Corporation/A Harcourt Assessment Company; Toronto, Canada: 2004.

Stavrakaki S. Comprehension of reversible relative clauses in specifically language impaired and normally developing greek children. Brain and Language 2001;77(3):419–431. [PubMed: 11386707]

Stowe LA. Parsing WH-constructions: evidence for on-line gap location. Language and Cognitive Processes 1986;1(3):227–245.

Swinney, D.; Ford, M.; Bresnan, J. On the temporal course of gap-filling and antecedent assignment during sentence comprehension.. In: Grosz, B.; Kaplan, R.; Macken, M.; Sag, I., editors. Language structure and processing. CSLI; Stanford, CA: 1989.

Van der Lely H. Jenkins L. Evidence for and implications of a domain-specific grammatical deficit. The genetics of language. 2004

Van der Lely H. Domain-specific cognitive systems: insight from Grammatical specific language impairment. Trends in Cognitive Sciences 2005;9(2):53–59. [PubMed: 15668097]

Van der Lely H, Battell J. Wh-movement in children with grammatical SLI: a test of the RDDR hypothesis. Language 2003;79(1):153–181.

van der Lely H, Rosen S, McClelland A. Evidence for a grammar-specific deficit in children. Current Biology 1998;8(23):1253–1258. [PubMed: 9822577]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

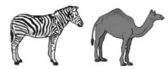

Fig. 1.
Primed and control picture probes for prime word “zebra”

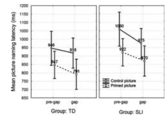

Fig. 2.

Mean reaction times per group, probe type and probe position in Experiment 2. *Error bars* represent 95% confidence intervals

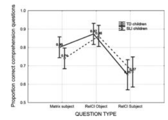

Fig. 3.

Mean accuracy on comprehension questions by subject group and question type in Experiment 2. *Error bars* represent 95% confidence intervals

Table 1

Bonferroni-protected comparisons, showing means and standard deviations per cell, and Bonferronicorrected *t*-test probabilities (pooled MSE = 48573, *df* = 30.79)

Group	*N*	Probetype	Probe position	Mean	SD	Priming score	Post-hoc *p*
SLI	14	control	Pre-gap	1060	230	–138*	0.016
SLI	14	primed	Pre-gap	922	151		
SLI	14	control	Gap	975	161	–105	0.14
SLI	14	primed	Gap	870	175		
TD	14	control	Pre-gap	946	123	–99	0.21
TD	14	primed	Pre-gap	847	143		
TD	14	control	Gap	918	163	–127*	0.029
TD	14	primed	Gap	791	151		

* marks significance