Published in final edited form as:

J Speech Lang Hear Res. 2013 April ; 56(2): 577–589. doi:10.1044/1092-4388(2012/11-0254).

Sentence Comprehension in Specific Language Impairment: A Task Designed to Distinguish between Cognitive Capacity and Syntactic Complexity

Laurence B. Leonard,
Purdue University, West Lafayette, IN

Patricia Deevy,
Purdue University, West Lafayette, IN

Marc E. Fey, and
University of Kansas Medical Center, Kansas City, KS

Shelley L. Bredin-Oja
University of Kansas Medical Center, Kansas City, KS

Abstract

Purpose—This study examined sentence comprehension in children with specific language impairment (SLI) in a manner designed to separate the contribution of cognitive capacity from the effects of syntactic structure.

Method—Nineteen children with SLI, 19 typically developing children matched for age (TD-A) and 19 younger typically developing children (TD-Y) matched according to sentence comprehension test scores responded to sentence comprehension items that varied in either length or their demands on cognitive capacity based on the nature of the foils competing with the target picture.

Results—The TD-A children were accurate across all item types. The SLI and TD-Y groups were less accurate than the TD-A group on items with greater length and, especially, on items with the greatest demands on cognitive capacity. The types of errors were consistent with failure to retain details of the sentence apart from syntactic structure.

Conclusions—The difficulty in the more demanding conditions seemed attributable to interference. Specifically, the children with SLI and the TD-Y children appeared to have difficulty retaining details of the target sentence when the information reflected in the foils closely resembled the information in the target sentence.

Deficits in sentence comprehension have been well documented in children with specific language impairment (SLI). Studies have shown that even children described as having an expressive language disorder perform below age level on tests of sentence comprehension (Bishop, 1979, 1982). Comprehension difficulties have been reported for a wide variety of sentence structures, including dative sentences (e.g., *Give the girl the boy*), passive sentences (e.g., *The cow is pushed by the boy*), relative clauses (e.g., *The dogs that are running are at the beach*), sentences with postmodified subjects (e.g., *The chicken on the ball is black*), wh-questions (e.g., *Who was the happy little girl washing?*), and sentences referring to continuous actions in the past (e.g., *Where was Minnie Mouse drawing the*

happy face?), among others (Bishop, 1982; Bishop & Adams, 1992; Bishop, Bright, James, Bishop, & van der Lely, 2000; van der Lely & Harris, 1990; Deevy & Leonard, 2004; Leonard & Deevy, 2010; Rispens & Been, 2007; Robertson & Joanisse, 2010; van der Lely & Stollwerk, 1997). These sentence comprehension difficulties seem surprisingly resistant to treatment (Bishop, Adams, & Rosen, 2006).

The nature of these sentence comprehension problems is a subject of considerable debate. Two schools of thought predominate in this discussion. According to one school of thought, particular grammatical principles are late to emerge, or are otherwise absent from the grammars of children with SLI, leading to misinterpretation of the related syntactic constructions (e.g., van der Lely, 2005).

A second school of thought holds that sentence comprehension weaknesses in children with SLI might be due to limitations in cognitive capacity. Although the types of grammatical comprehension difficulties attributed to children with SLI are not in question, those favoring a cognitive capacity limitation account do not find sufficient evidence to conclude that the difficulties are due to a disruption of particular grammatical principles (Bishop et al., 2000; Norbury, Bishop, & Briscoe, 2002).

Many well-designed studies have reported relationships between sentence comprehension problems and capacity limitations. However, to date, the studies have relied on procedures that are either indirect measures of this relationship, or do not succeed in separating the possible effects of poor understanding of grammatical structure from the cognitive capacity demands that are experimentally manipulated. The goal of the present study was to examine the role of cognitive capacity in the sentence comprehension of children with SLI in a way that significantly reduces any possible confound of difficulties with the grammatical structure itself.

Examining the Role of Cognitive Capacity in Sentence Comprehension

Cognitive capacity refers to the amount of computational space or energy needed to perform a mental task (Kail & Salthouse, 1994). Although the term might suggest a passive repository of information, capacity is actually a broad ability that can involve a wide range of cognitive processes including attention, speed of processing, the use of processing strategies, as well as the storage, rehearsal, and retrieval of information, among others. A growing literature has provided evidence that children with SLI have weaknesses in a variety of these areas. For example, recent studies show weaknesses in tasks of sustained attention in these children (e.g., Finneran, Francis, & Leonard, 2009; Spaulding, Plante, & Vance, 2008). Slow processing speed has also been well documented in children with SLI, with reduced speed observed during nonlinguistic as well as linguistic tasks (e.g., Kail, 1994; Miller, Kail, Leonard, & Tomblin, 2001; Leonard et al., 2007; Windsor & Hwang, 1999). Weaknesses in storage have been seen through tasks of nonword repetition that tap phonological short-term memory (see meta-analysis by Graf-Estes, Evans, & Else-Quest, 2007). Other investigators have employed measures of working memory, that is, measures designed to capture both storage and mental manipulation. A commonly used task of this type is the listening span task, in which children must listen to sentences and respond to their truth value, while also retaining the last word in each sentence for subsequent recall. Children with SLI show weaknesses in this type of activity (e.g., Archibald & Gathercole, 2006; Ellis Weismer, Evans, & Hesketh, 1999; Mainela-Arnold & Evans, 2005; Marton & Schwartz, 2003). Tasks involving visual-spatial memory also provide evidence of weakness in these children (Bavin, Wilson, Maruff, & SLeeman, 2005; Hoffman & Gillam, 2004).

It is not difficult to see how weaknesses in these and other cognitive processes could hinder children's language development. To acquire language, children must build up lexical and

grammatical representations based on information in the input. However, for children with limitations in attention, processing speed, information storage, rehearsal, retrieval, or related processes, comprehension of the language in the input would be only partial, and therefore lexical and grammatical representations stored in long-term memory would be built up only slowly and, possibly, incorrectly.

Many of the studies examining the relationship between cognitive capacity and sentence comprehension have employed separate tasks for each and have then determined the correlation between them. Significant positive correlations have been found between nonword repetition and sentence comprehension as well as between listening span and sentence comprehension (e.g., Montgomery, 1995, 2000; Norbury et al., 2002; Robertson & Joanisse, 2010). However, the strength of the relationship may depend on the types of sentences used in the comprehension task. For example, Montgomery and Evans (2009) found that nonword repetition correlated with simple sentence comprehension in children with SLI, but listening span correlated with the comprehension of complex sentences.

Investigators have also manipulated the types of sentences used in comprehension tasks with the intent of gaining insight into the contribution of working memory. For example, Montgomery (1995, 2000) found that children with SLI performed below the level of younger typically developing peers on sentences containing “redundant” elements but did not differ from their peers on sentences with the redundant elements removed. The sentences with redundant elements were always longer than the non-redundant sentences and differed further from the non-redundant sentences either structurally (e.g., *The girl who is crying is pushing the boy who is smiling* versus *The girl crying is pushing the boy smiling*) or in the number of modifying adjectives in the sentence (e.g., *The dirty little boy climbs the great big tall green tree* versus *The little boy climbs the tree*).

Deevy and Leonard (2004) examined wh-question comprehension in children with SLI and typically developing children matched on receptive vocabulary scores. The two groups were similar and quite accurate in their comprehension of short wh-object questions such as *Who is the dog washing?* However, when the wh-object questions were lengthened by adding superfluous adjectives, as in *Who is the happy brown dog washing?*, the accuracy of the children with SLI dropped significantly and fell below that of the vocabulary-score control children. Deevy and Leonard argued that this drop in performance appeared to be related to retention factors, because the syntactic operations involved in short and long wh-object questions (relating the wh-word to the object position that follows the verb) are the same. In the case of long wh-object questions, the children had to retain the wh-word for a longer distance (and hence for a longer time) before its interpretable position was encountered, given the intervening adjectives. These retention demands seemed to be a larger obstacle for the children with SLI than for the control children.

Robertson and Joanisse (2010) manipulated cognitive load in sentence comprehension in three ways. First, sentences had either canonical (for English) subject-verb-object word order (e.g., *This is the girl that pinches the doctor*) or noncanonical word order (e.g., *This is the man that is pointed at by the boy*). Second, sentences were either relatively short (as in the preceding examples) or long (e.g., *This is the boy in the bright red pants that taps the girl with the nice blond hair*). Third, the timing between the presentation of the pictorial stimuli and the auditory presentation of the test sentence was varied (e.g., the picture appearing 2 sec before the sentence began; the picture appearing 3 sec after the sentence was completed). The children with SLI were less accurate than same-age controls and a group of children with dyslexia, but did not differ from a group of younger controls matched on receptive vocabulary scores. Each of the three types of manipulations had an effect on the children’s performance. However, the nature of the errors by the children with SLI pointed to syntactic

problems even when the cognitive loads were minimal. When cognitive demands were increased, performance dropped further.

The sentence length manipulation used by Robertson and Joanisse (2010), however, represented more than the simple addition of length. The stimulus sentences in the “long” condition added post-modifying prepositional phrases to the subject and object noun phrases included in each of the “short” sentences (e.g., *This is the boy in the bright red pants that taps the girl with the nice blond hair*). The action depicted was the same for all four pictures in the array (e.g., one character tapping on the shoulder of another character). The target picture and one of the foils were distinguishable only by an attribute described by the adjective in either the subject or object noun phrase (e.g., the boy doing the tapping had blue pants rather than red pants). Thus, to respond accurately to the long sentence in this example, children had to comprehend the sentence and retain the critical descriptive information [Boy-RedPants-ActsOn-Girl] while searching to identify the target picture. Because only the attribute of the subject or the attribute of the object was manipulated, it was sufficient for the second noun (*girl*, in this example) to be retained without an attribute. This manipulation clearly increased comprehension and cognitive demands relative to short sentences such as *This is the girl that pinches the doctor*, because in the latter the children had to comprehend the sentence and retain only [Girl-ActsOn-Doctor]. However, because cognitive load increased along with length, the source of the effect is not clear. For example, there may have been an effect of adding attributes, such as *with the nice blond hair*, even though this information was not essential to form a correct response (as all girls in the array had blond hair).

In the present study, we adopted a somewhat different approach to enable us to better distinguish the effects of syntactic structure and cognitive capacity. All three item types in our task required the same type of response – finding the picture out of an array of four that reflected the relationship conveyed in the sentence. Our “low demand” items required interpretation of basic subject-verb-object relations (e.g., *The dog washes the pig*). Our “intermediate demand” items added length by attaching adjectives to the subject and object nouns in the low demand sentences. These adjectives were semantically superfluous, however. For example, for the item, *The happy dog washes the little pig*, the dogs and pigs that appeared in the foils were identical to the dog and pig shown in the target picture; the attributes were not contrastive, as all dogs were happy and all pigs were little. Our “high demand” items added significant cognitive capacity demands, even though the length and syntactic structure of these items were the same as in intermediate demand items. An example is the item, *The yellow dog washes the white pig* (see Figure 1). In this condition, the information introduced by each adjective was contrastive. For example, one foil depicted a brown dog washing a white pig, and another showed a yellow dog washing a pink pig. Thus, to respond correctly, the children had not only to understand the proper subject-verb-object relationship, but also to remember the particular attribute associated with each of the characters as they searched for the correct drawing.

The burden placed on cognitive capacity in the high demand condition can be attributed to the fact that the children had to retain the details of the target sentence in the face of information from the foils that closely resembled that of the target sentence. The similarity between the target sentence and the information from the foils created the potential for interference. Thus, any difficulty with high demand items cannot be attributed to verbal storage in isolation. Consider the difference between sentences in the intermediate demand condition (containing superfluous adjectives) such as *The happy dog washes the little pig* and those of the high demand condition (containing contrastive adjectives) such as *The yellow dog washes the white pig*. If considered solely in terms of verbal storage (e.g., if presented in a sentence repetition task with no pictures), these two sentences are highly

similar. However, in the case of the high demand item, as children search through the four alternative drawings, they see drawings depicting a dog with the correct attribute washing a pig with the wrong attribute, a dog with the wrong attribute washing a pig with the correct attribute, as well as a dog and pig with the correct attributes but with incorrect subject-object roles. Thus the challenge to retention of the target sentence comes from competing information.

Note that although the greater potential for interference renders the high demand condition more challenging, we cannot attribute the vulnerability to interference to any one cognitive process. For example, weaknesses in attention or response inhibition could allow children to select a similar but incorrect drawing, reduced speed of processing could limit the time available for rehearsal of the target sentence, cross-modal (verbal and visual) information may be poorly coordinated, or other weaknesses singly or in combination could lead children to be vulnerable to the greater potential for interference that is inherent in the high demand items. For this reason, we employ the more general term “cognitive capacity” in favor of a term that implies greater specificity than our task warrants. An important advantage of our task is that by comparing sentences whose adjectives are either contrastive or superfluous, we are able to manipulate cognitive capacity demands while controlling for both length and syntax.

We ensured that the children with SLI exhibited good comprehension of the syntactic structure to be used as well as the particular adjectives to be included. As a result, we could compare the children’s performance with that of their typically developing peers when cognitive capacity demands increased through manipulation of the foils that competed with the correct drawing. We predicted that as capacity demands increased, the performance of the children with SLI would show greater declines than that of their age-matched peers. We also recruited a group of younger typically developing children who were matched with the children with SLI on a test of sentence comprehension. Previous studies have yielded mixed results, with some studies showing greater accuracy by younger controls than by children with SLI when processing demands were placed on syntactic comprehension (Deevy & Leonard, 2004; Montgomery, 1995), whereas other studies have found comparable accuracy levels for these two groups (Robertson & Joanisse, 2010). Given the systematic manipulation of cognitive capacity demands across the three conditions in the present study, we hoped to detect the point at which these two groups differed, thus giving us a clearer picture of the status of cognitive capacity in children with SLI when comprehension of syntactic structure is controlled.

Method

Participants

Fifty-seven children participated in the study, with 19 children in each of three groups. Descriptive data are provided in Table 1. One group of children formed the SLI group. All 19 of these children had been diagnosed with language impairment and were participating in group or individual language intervention programs. They ranged in age from 48 to 66 months. All children met the following inclusionary criteria: They passed a hearing screening, an oral motor screening, and showed no symptoms of disturbance in relating to persons or objects. In addition, each child scored above 85 on the *Columbia Mental Maturity Scale* (CMMS, Burgemeister, Blum, & Lorge, 1972). This measure served only to ensure adequate nonverbal functioning; it was not used for matching children across groups. A two-step process was used to document these children’s language difficulties. First, we determined if the children met the cutoff score of 87 on the *Structured Photographic Expressive Language Test – Preschool 2* (SPELT-P2, Dawson, Eyer, & Fonkalsrud, 2005) determined by Greenslade, Plante, and Vance (2008) to be the cutoff point yielding high

sensitivity and specificity for this age group. Second, If any children scored just above this cutoff score, they were included if their Developmental Sentence Score (Lee, 1974) was below the 10th percentile. Eighteen of the 19 children selected for the SLI group met the first criterion; the remaining child scored 89 on the SPELT-P2 but earned a DSS that fell below the 10th percentile. The children were also administered the Sentence Structure subtest of the *Clinical Evaluation of Language Fundamentals – Preschool 2* (CELF-P2, Wiig, Secord, & Semel, 2004). However, this measure was used to match these children with younger typically developing children (see below), and was not used as a basis for inclusion.

A second group of 19 children served as an age-matched comparison group. Each child in this group was within two months of the age of a child in the SLI group. They ranged in age from 48 to 68 months. These children passed the same screening tests and tests of nonverbal functioning that were administered to the children with SLI. Their SPELT-P2 scores were 92 and above. Although CELF-P2 scores were not used for selection purposes, it can be seen from Table 1 that these children's CELF-P2 scores were considerably higher than those of the children with SLI, $t(36) = 4.98$, $p < .001$, $d = 1.62$. Hereafter, these age-matched typically developing children are referred to as the TD-A group.

The 19 children in the third group were younger, typically developing children, ranging in age from 37 to 47 months. Each child in this group earned a raw score on the Sentence Structure subtest of the CELF-P2 that was within 2 points of the score of a child in SLI group. On average, these children were approximately 15 months younger than the children in the SLI group. Their scores on the SPELT-P2 were 91 and above. These children passed the same screening tests and tests of nonverbal functioning described above. However, the CMMS could not be used with nine children below 42 months of age due to the age range of the standardization of this test. For children in this group below 42 months of age, age-appropriate nonverbal functioning was demonstrated on the *Leiter International Performance Scale – Revised* (LIPS, Roid & Miller, 1997). Hereafter, we refer to the 19 children in this group as the TD-Y group.

Although all children scored within normal limits in nonverbal functioning, the three participant groups were not equivalent in this regard. Recall that nine of the 19 TD-Y children received a different nonverbal test (given their younger ages) than all remaining children. Therefore, to compare non-verbal cognitive function across groups pre-experimentally, we converted the children's standard scores to age scores, using the formula, $MA = \text{Standard Score} / 100 * CA$. A comparison based on analysis of variance revealed a significant difference among the groups, $F(2, 54) = 30.96$, $p < .001$, partial $\eta^2 = 0.53$. Post-hoc (Fisher LSD) testing revealed that the TD-A children's age scores ($M = 67.30$, $SD = 9.30$) were significantly higher than those of the children with SLI ($M = 59.35$, $SD = 8.56$, $p < .001$, $d = 0.89$) which, in turn, were significantly higher than the age scores of the TD-Y children ($M = 46.96$, $SD = 5.82$, $p < .001$, $d = 1.69$).

Procedure

Children were initially administered three sets of screening items designed to ensure that they comprehended the vocabulary and syntactic structures to be used in the experimental task. One set consisted of 10 items designed to assess the children's understanding of simple reversible subject-verb-object sentences involving the same nouns and verbs used in the experiment proper, as in *The horse kicks the cow*. Each item involved a spoken sentence digitally recorded by a female speaker and two drawings on a computer screen. One drawing corresponded to the test sentence (e.g., a horse kicking a cow), the other to the reverse (a cow kicking a horse). Children were seen individually. The children were told that they would see some pictures on the computer and were to point to the picture that matched what "the lady says." Each item proceeded in the following way. When the child was attentive,

the experimenter activated the program so that the two drawings appeared on the screen. A tone was then presented followed 1.5 sec later by the audio presentation of the test sentence. The child responded by pointing to the picture that matched the sentence. The experimenter then activated the program to show the next set of drawings, and the procedure continued as described for the first item. The experimenter had manual control of the pace of item presentation in case a child became momentarily distracted and it was necessary to delay the next item until the child's attention was re-directed to the task.

A second set of screening items consisted of 20 items that assessed the children's understanding of the adjectives to be used in the experimental task. For these items, the test sentence took the form seen in *Show me the red bird*. In all items, the adjective was tested in pre-nominal position (e.g., red bird, happy turtle, little frog). One of the two drawings depicted the attribute expressed by the adjective, the other drawing showed an attribute that was quite different from the one mentioned in the test sentence (e.g., a blue bird, a sad turtle, a large frog).

The third set of screening items was composed of 10 items involving the reversible subject-verb-object structure with adjectives preceding both the subject and object name, as in *The brown monkey feeds the little tiger*. Again, the child had to select from two drawings on the screen. The foil depicted the opposite of the relationship mentioned in the target sentence (a tiger feeding a monkey). Importantly, the adjectives were irrelevant to the task, as the characters shown in the alternative drawing had the same attributes as those in the target drawing (e.g., both monkeys were brown, both tigers were little).

Table 2 provides a summary of the children's performance on the three sets of screening items. As can be seen from this table, the three groups performed at very similar levels. All children responded correctly to at least 70% of the items on each of the three sets of screening items. The lowest performance was seen by one child in the TD-Y group who scored 70% on both the simple subject-verb-object items and the subject-verb-object items with superfluous adjectives, and 90% on the adjective items. Together, these sets of screening items provided a good basis for assuming that the children had basic knowledge of the linguistic material to be used in the experimental task. Specifically, they understood the adjectives, the subject-verb-object structure, and could also succeed even with the added length created by the inclusion of (superfluous) adjectives. Because these items contained only two response options, they placed little burden on cognitive capacity.

The experimental task was also divided into three sets of items, administered in counterbalanced order on separate days. The experimental items differed from the screening items in that four, rather than two drawings appeared on the computer screen that the children had to select from. All four drawings portrayed the same action being performed by a subject on an object. In other respects the procedure was identical to that used for the screening items. Ten items were used for each set. The sentences used for these items appear in the Appendix. One set of items consisted of simple reversible subject-verb-object structures (e.g., *The bunny chases the cat*). Along with the target drawing, the array included a drawing depicting the opposite relationship (e.g., a cat chasing a bunny), a drawing showing the correct subject acting on an object not named in the test sentence (e.g., a bunny chasing a chicken), and a drawing showing a subject not named in the test sentence acting on the correct object (e.g., a chicken chasing a cat). It can be seen, then, that one foil represented a syntactic reversal error, and the other two portrayed either the correct subject-verb relationship or the correct verb-object relationship but the wrong character was represented (e.g., a chicken) for the remaining constituent. Although the four alternatives rendered these subject-verb-object items somewhat more difficult than the two-alternative items used for screening, relative to the remaining sets of items in the experimental task (see

below), these items were not expected to be especially difficult. They required only comprehension and retention of the type [Bunny-ActsOn-Cat]. Accordingly, we refer to them here as constituting the “low demand” condition.

Another set of 10 items consisted of reversible subject-verb-object sentences containing pre-nominal adjectives, as in *The nice mouse covers the pretty bird*. The adjectives were superfluous; they added to the length of the test sentences but were not contrastive. The foils for these items were similar to those used in the preceding set. One foil depicted the opposite relationship (e.g., a bird covering a mouse), one showed the correct subject acting on an unnamed object (e.g., a mouse covering a frog), and the remaining foil showed an un-named subject acting on the correct object (e.g., a frog covering a bird). Two features of the drawings prevented the children from treating the adjectives as contrastive. First, when a character appearing in the target sentence (e.g., a nice mouse, a pretty bird) was also used in a foil, the drawings of the character were made to render them as much the same as possible. Second, only one drawing – the target picture – portrayed both characters in the proper syntactic relationship (e.g., only the target picture showed a mouse covering a bird). Consider a case in which a child determines that the mouse in the target picture does not look “nice” (even though the mice in all of the drawings were essentially identical). For this reason, the child thinks about choosing an alternative to the correct drawing. Because of the options available, the child would have to choose an alternative drawing that showed the wrong syntactic relationship or an alternative that depicted characters not even mentioned in the target sentence. Therefore, the target picture was always clearly the best response. Although the sentences in this set required only knowledge of the basic subject-verb-object structure and retention of the character names (as in [Mouse-ActsOn-Cat]) to arrive at a correct response, the added length created by the inclusion of the superfluous adjectives should have rendered retention of relevant details of the sentence somewhat more difficult. For this reason, we refer to these items as constituting the “intermediate demand” condition.

Finally, 10 items were used in what we consider to be the “high demand” condition. These items, too, employed pre-nominal adjectives in reversible subject-verb-object structures, as in *The yellow dog washes the white pig*. However, each adjective in the test sentence was contrastive. One foil depicted the opposite relationship (e.g., a white pig washing a yellow dog). Another foil showed the named subject with an incorrect attribute acting on the correct object with the correct attribute (e.g., a brown dog washing a white pig). The remaining foil portrayed the correct subject with the correct attribute acting on the named object with an incorrect attribute (e.g., a yellow dog washing a pink pig). To respond correctly to such items, the children had to comprehend the sentence and retain information of the type [Yellow-Dog-ActsOn-White-Pig] as they searched for the correct drawing; failure to retain any portion of this information could result in an incorrect choice. The four drawings used for one of the items in the high demand condition can be seen in Figure 1.

The children’s responses were scored for accuracy. In addition, their incorrect selections were noted for subsequent error analysis.

Results

The children’s accuracy levels were examined through a repeated measures analysis of variance with participant group (SLI, TD-A, TD-Y) as a between-subjects variable and demand condition (low demand, intermediate demand, high demand) as the within-subjects variable. Largely because of near ceiling performances of the children in the TD-A group, the distributions were generally not normal and the data failed to meet assumptions of sphericity and homogeneity of variance. Therefore, all data were arcsine transformed before analysis. This step rendered the data suitable for repeated measures analyses. To facilitate

clarity and interpretation, Figure 2 presents untransformed means and 95% confidence intervals. Means and standard deviations for the untransformed scores and the arcsine transformations are reported in Table 3.

Main effects were found for participant group, $F(2, 54) = 13.65, p < .001$, partial $\eta^2 = .34$ and for demand condition, $F(2, 108) = 14.71, p < .001$, partial $\eta^2 = .21$. However, a marginally significant participant group $\times$ demand condition interaction was also apparent, $F(4, 108) = 1.98, p = .10$, partial $\eta^2 = .07$. Planned comparisons were performed, focusing on whether, as expected, the accuracy of the children with SLI would decline on high demand items relative to intermediate demand items. This comparison was especially important given that items in these two conditions shared length, syntactic structure, and vocabulary. We also tested our prediction that differences between the children with SLI and their same-age peers would increase in magnitude as demand increased, from intermediate to high demand items. The SLI group's scores were indeed lower in the high demand condition than in the intermediate demand condition ($p = .02, d = 0.71$). It was also the case that the differences between the SLI group and the TD-A group increased in magnitude from the intermediate demand condition ($p = .002, d = 1.12$) to the high demand condition ($p < .001, d = 1.82$). In contrast to the findings for the comparison between the intermediate and high demand condition, we found no difference in the SLI group's performance on the low and intermediate demand items ($p = .38, d = 0.21$). The TD-Y children, like the children with SLI, were less accurate on items in the high demand condition than on items in the intermediate demand condition ($p = .05, d = 0.48$). However, unlike the children with SLI, the TD-Y children found the intermediate demand items more difficult than the low demand items ($p = .02, d = 0.52$). As can be seen from Figure 2, this difference from the SLI group stemmed from non-significantly stronger performance by the TD-Y group on the low demand task. The children with SLI and TD-Y performed similarly on the intermediate task.

It is important to note that even in the most challenging condition (high demand), both the children with SLI and the TD-Y children's performance was above chance. With four alternative drawings to select from, chance level is 25%, yet the SLI and TD-Y groups showed accuracy levels of 59.47% and 60.00%, respectively.

The findings for the comparison between the SLI and TD-Y groups are somewhat difficult to interpret, because the children in the TD-Y group had lower age scores than the children in the SLI group. To further complicate the situation, for the TD-Y group, these age scores were estimated from two different tests, the CMMS and the LIPS, which differ somewhat in the cognitive processes they assess. To arrive at a clearer interpretation of the performance patterns of these two groups, then, we needed to control cognitive ability to a greater degree.

Accordingly, in a second analysis, we excluded the nine TD-Y participants who took the LIPS, thus ensuring that the measure of nonverbal cognition was the same for all participants, the Maturity Index from the CMMS (see Table 1). This step had the effect of eliminating the statistical difference in nonverbal intelligence between the TD-Y and SLI groups, $t(27) = 1.27, p = .21$. Importantly, these two groups were also highly similar in their CELF-P2 Sentence Structure scores, $t(27) = 0.43, p = .67$. We then compared the two groups through a repeated measures analysis of variance, with demand condition as the within-subjects variable.

These attempts to control cognitive ability as measured by the CMMS did not change the results of our analyses substantially. As in the original analysis, the SLI and TD-Y groups performed at similar levels, $F(1, 27) = 0.41, p = .53$, partial $\eta^2 = .02$. As before, a main effect for demand condition was observed, $F(2, 54) = 13.49, p < .001$, partial $\eta^2 = .33$. Planned comparisons revealed that scores were lower in the high demand than in the

intermediate demand condition ($p = .02$). In addition, scores for the intermediate demand condition items were lower than those for the low demand condition items ($p = .02$). Recall that in the first analysis, the children with SLI scored at statistically comparable levels on items in the low and intermediate demand condition. The SLI data entered in this second analysis were the same as in the first analysis, as we retained all 19 children with SLI. The difference between low and intermediate items in the second analysis, then, suggests that this subset of TD-Y children, like the larger TD-Y group of the first analysis, found the low demand items to be significantly easier than the intermediate demand items. Nevertheless, the participant group $\times$ demand condition was not significant, $F(2, 54) = 0.80$, $p = .45$, partial $\eta^2 = .03$.

The children's errors were examined next to determine if they would shed light on the nature of the children's misinterpretations. The errors of all children (19 children in each of the three participant groups) were included for this purpose. Table 4 provides a breakdown of the distribution of errors. Reversal errors were instances in which the children selected a drawing that depicted the appropriate characters and attributes but reversed their subject and object roles. For the low and intermediate demand conditions, incorrect object and incorrect subject errors refer to cases in which the children selected a drawing portraying a character in object or subject position that was not named in the target sentence. For the high demand condition, these errors refer to instances in which the child chose the character named in object or subject position but the particular drawing selected had the wrong attribute for the character (e.g., for the sentence *The yellow dog washes the white pig*, a drawing of a yellow dog washing a pink pig was selected).

The most obvious detail in Table 4 is the fact that reversal errors were the most frequent type of error for the low and intermediate demand conditions, but not for the high demand condition. This pattern held true for all three participant groups, even though the TD-A children made fewer errors overall. A second noteworthy finding from Table 4 is that, along with the dramatic increase in the number of choices of the incorrect object and incorrect subject in the high demand condition, there was a decrease in the absolute number of reversal errors in this condition. This was seen most clearly in the data from the children with SLI, although it is generally true for the other two participant groups as well.

To assess the statistical reliability of this trend, we compared the percentage of errors that constituted errors of reversal in the low demand and intermediate demand condition combined with the percentage of errors that represented errors of reversal in the high demand condition. Given that accuracy was relatively high in the low demand condition, it was necessary to combine the error data from the low and intermediate demand conditions to have sufficient data for analysis. One child each from the SLI and TD-Y groups and nine children from the TD-A group made no errors in the low + intermediate demand conditions and/or the high demand condition and were therefore excluded from consideration, resulting in an N of 18 children with SLI, 18 TD-Y children, and 10 TD-A children serving as the source of the data. An analysis of variance with condition (low + intermediate versus high) and participant group (SLI, TD-Y, TD-A) was calculated. The dependent variable was the percentage of total errors represented by each error type; percentage data were arcsine transformed to stabilize variance. A significant main effect was found for condition, $F(1, 43) = 27.99$, $p < .001$, partial $\eta^2 = .39$, revealing a much larger percentage of errors constituting errors of reversal in the low + intermediate demand conditions ($M = 56.39\%$, $SD = 33.60$) than in the high demand condition ($M = 18.95\%$, $SD = 23.55$). No main effect was found for participant group, $F(2, 43) = 0.08$, $p = .92$, $\eta^2 = .01$, and no interaction was apparent, $F(2, 43) = 0.82$, $p = .45$, $\eta^2 = .04$.

Finally, there seemed to be a tendency for the children with SLI to make more errors of choosing the incorrect object than choosing the incorrect subject in the low demand condition, whereas the TD-Y children showed the reverse trend. These differences seemed to characterize the groups as a whole. Unfortunately, given the generally low frequency of these errors (as can be seen in Table 4), we were unable to test the statistical reliability of this trend.

Discussion

In this study, the children with SLI showed the same level of accuracy as younger children matched on a test of sentence comprehension (the TD-Y children), while performing significantly below the level of same-age peers (the TD-A children). The high demand condition was shown to be more difficult than the remaining conditions. Of special importance was the finding that the children with SLI were less accurate on items in the high demand condition than on items in the intermediate demand condition. The distribution of errors differed across the demand conditions. Reversal errors were much more likely in the low and intermediate demand conditions than in the high demand condition. For the high demand condition – for which there were more errors overall – selections of foils with the wrong attribute-subject pairing or wrong attribute-object pairing were more numerous than errors of reversal.

Together these findings lead us to two major conclusions: (1) the experimental task seemed to separate effects relating to cognitive capacity from those relating to syntactic comprehension; and (2) interference resulting from similarity between the target sentence and information reflected in the foils contributed to differences among the conditions as well as to differences among the groups of children. We first discuss these conclusions in terms of what they say about cognitive capacity limitations in the children with SLI. We then turn to a discussion of how best to interpret the similar performance levels of the SLI and TD-Y groups.

The Nature of the Cognitive Capacity Limitations in the Children with SLI

Two observations suggest that cognitive capacity, rather than syntactic comprehension, played a key role in contributing to the weak performance of the children with SLI relative to the TD-A children. First, the children with SLI, like the TD-A children, showed high performance levels on the screening items, indicating good comprehension of the lexical content and syntax employed in the experimental task. Second, although the sentences in the intermediate and high demand conditions were essentially identical in length, syntactic structure, and lexical content, the children with SLI were significantly less accurate on high demand items than on intermediate demand items whereas the TD-A children were accurate on both types of items. Judging from the effect sizes, the difference between these two groups was considerably larger in the high demand condition than in the intermediate demand condition.

Although these observations seem to point to cognitive capacity limitations as an important factor in distinguishing the children with SLI from the TD-A children, our task was not designed to isolate the specific processes that might have been most responsible for these limitations. Several different processes are viable candidates. For example, the similarities of the drawings in the high demand condition made attention an especially important process. Processing speed would have been beneficial to ensure time for rehearsal after the information had been stored. Furthermore, children had to show inhibitory control to refrain from responding to the first foil encountered that bore some resemblance to the target sentence.

Although the precise sources of weakness are not clear, it does appear that these weaknesses led to a vulnerability to interference. The potential role of interference can be seen by considering the requirements of the task. Appropriate responding to each target sentence required storage of verbal information as the sentence was interpreted, based on knowledge stored in long-term memory. This information then had to be retained while the child searched through the array of four alternative drawings, comparing each drawing to information in the target sentence. We assume that the children performed this comparison process in either of two general ways (or some combination of both). The children may have converted visual information (details in the drawing) to verbal information (e.g., “yellow dog”, “white pig”) and compared it to the stored verbal details of the target sentence. Alternatively, the children may have directly compared the target sentences to the visual information in each drawing.

The only difference between the sentences in the intermediate and high demand conditions is that in the high demand condition, the information in all three foil drawings bore considerable similarity to the target sentence. For example, in one picture, each adjective was paired with the proper character but the roles of the characters were reversed. In the other two foils, both the characters and the roles were correct but either the subject or object attribute did not match that of the target sentence. Thus, there was a greater amount of potentially interfering information in the foils of the high demand condition. This interference, we believe, placed far greater demands than did the other conditions on attention to details of the drawings and sentences, on retention of the details of the target sentence, and on inhibition of responses to drawings that differed from the information in the target sentences only in small ways.

According to this interpretation, the observed pattern of errors occurred because in the intermediate condition, drawings that were the most similar to the target sentence depicted the correct characters in a reversed relationship. In contrast, in the high demand condition, drawings that were the most similar to the target sentence were those that depicted the correct characters in the correct relationship but with one incorrect attribute. We propose further that there were fewer errors overall in the intermediate condition because there was less opportunity for interference than in the high demand condition.

One workable alternative explanation for the difference in accuracy between the intermediate and high demand conditions is that the children with SLI retained only partial details of the target sentence from the outset, such as the character names and their respective subject/object roles. In this case, the children would have searched for a drawing that corresponded to, say, *The dog washes the pig* (instead of *The happy/yellow dog washes the little/white pig*). Because two of the foils in the intermediate condition depicted unnamed characters and could easily be eliminated as options and because the adjectival information was not needed in the intermediate condition, this limited information would have been enough to preserve relatively high rates of accurate responding in the intermediate task. In contrast, in the high demand condition, there were no foils that showed unnamed characters, and information represented by the adjectives was crucial to correct responding. Thus, any incompletely stored form would be expected to lead to more errors in the high demand condition than in the intermediate demand condition, as was the case.

One weakness of this alternative account is that it does not provide a clear reason for why most errors in the intermediate condition were errors of reversal, whereas those in the high demand condition were selections of the correct subject-object relationship but with an incorrect attribute associated with either the subject or object. If the children's errors were due to an incomplete stored form, the percentage of errors constituting errors of reversal should have been similar across all conditions. In fact, not only was the percentage of errors

involving reversal higher in the intermediate demand condition (61%) than in the high demand condition (16%), but the frequency of such errors was higher as well (27 versus 12), even though many more errors occurred overall in the high demand condition than in the intermediate demand condition. We believe that our interpretation provides a more straightforward reason for the observed pattern of errors: The most frequent error type in each condition was the selection of a drawing that bore the closest resemblance to the information in the target sentence. Hence, it had the most potential for interference.

Another alternative interpretation is that decay, rather than interference, was a major factor in the relatively weak performance of the children with SLI in the high demand condition. However, a decay interpretation does not seem adequate to explain the performance difference between the intermediate and high demand condition for these children. Both conditions required the children to retain sentences of the same length and structure and scan through four alternative drawings. It would seem that a decay interpretation would have to assume that decay is more likely for certain types of information – in this case, adjectives – than for other types of information. However, if this were true, selections of the drawing that depicted the wrong adjective with the object would seem less likely than selections showing the wrong adjective with the subject, due to the benefits of recency effects. That is, adjectives appearing later in the sentence would have less opportunity to decay than adjectives appearing earlier. Yet these two types of errors appeared with similar frequency in the responses of the children with SLI.

If our assumptions about interference effects are correct, the findings of the present study have implications for clinical assessment. In particular, it may not be safe to assume that tests of sentence comprehension are measuring only a child's understanding of syntactic structure, even in such "simple" tasks as picture pointing. In the present study, we documented that the children in the study comprehended subject-verb-object sentences with and without modifying adjectives. Yet we saw how accuracy levels changed dramatically as a function of the types of foils used along with the target drawing.

How Closely Did the Children with SLI Resemble the TD-Y Children?

Although the SLI and TD-Y groups performed in a similar manner based on our first analysis, certain details distinguished these two groups. When we converted the children's nonverbal test scores to age scores, we found that the age scores of the TD-Y children were lower than those of the children with SLI. The TD-Y children also differed from the children with SLI in their performance profile on the sentence comprehension task. Whereas the scores of children with SLI were similar for the low and intermediate demand items, the TD-Y children were less accurate on intermediate demand items than on the low demand items.

It seemed possible that the relative immaturity of the TD-Y children as reflected in their lower age scores contributed to this difference between these two groups. However, when we performed the second analysis, retaining only those TD-Y children who had received the CMMS, the SLI and TD-Y groups did not differ in their overall accuracy levels, but subtle differences between these two groups continued to hold. Again, it appeared that the TD-Y children had more difficulty with the intermediate demand items relative to the low demand items, a pattern that did not characterize the SLI group.

These subtle differences seem especially important considering that, in both analyses, the SLI and TD-Y groups were similar in their CELF-P2 Sentence Structure scores and, in the second analysis, were similar as well in their CMMS Maturity Index scores. As can be seen in Table 1, the SLI and TD-Y groups did not overlap in age and therefore, we cannot exclude some factor related to maturity as contributing to the differences between these groups, even in the second analysis. For example, the TD-Y children's younger ages might

have made them less proficient in handling the greater length seen in the intermediate demand sentences even though the adjectives in these sentences were superfluous. From Figure 2, it can be seen that the TD-Y children were non-significantly more accurate than the children with SLI on low demand items. The fact that these two groups were indistinguishable on intermediate demand items suggests that the TD-Y children may have been more adversely affected than the children with SLI by the additional material in the intermediate demand sentences.

It is also possible that the maturity differences of the SLI and TD-Y groups led to strategy differences that were not detectable from the accuracy scores alone. For example, although the numbers were not large, the TD-Y children were less likely than the children with SLI to retain subject information in the low demand condition. This might be expected if, due to their young age, the TD-Y children made less use of rehearsal. Without rehearsal, subject information would be more vulnerable to decay than object information, due to the earlier appearance of subjects in the sentence.

One important implication from these findings is that even when two groups seem closely matched on a measure of sentence structure comprehension, maturity-related processes might well lead to subtle differences in how the children approach a sentence comprehension task. These subtle differences might appear alongside the two groups' shared response pattern of decreased accuracy when capacity demands increase, as seen here in the SLI and TD-Y groups' significant decline in accuracy from the intermediate demand condition to the high demand condition.

Future Applications of the Sentence Comprehension Task

The advantage of the experimental task used in this study is that it allowed us to evaluate the role of cognitive capacity within sentence comprehension itself without altering the nature of the verbal material across conditions. In particular, the intermediate and high demand sentences were very similar, differing in their demands on the children only by virtue of the types of foils that competed with the target picture. Using this task, we found clear group differences that increased in magnitude as cognitive capacity demands increased.

Although this task seems to be successful in separating cognitive capacity from syntactic comprehension effects, we do not yet know if the effects of capacity demands operate in the same way regardless of how well the children have control of the syntactic structures employed in the target sentences. For example, if a syntactic structure is only weakly established in the children's grammars, a high demand condition might cause even greater declines in performance than observed in the present study. Such declines might exceed those seen in younger typically developing children.

Given this possibility, it is informative to revisit the investigation conducted by Robertson and Joannis (2010). The foils used by these investigators were similar to ours, though a correct response in their more challenging long sentence condition required retention of four elements (as in [Boy-RedPants-ActsOn-Girl]) rather than five elements in our own study (as in [Yellow-Dog-ActsOn-White-Pig]). However, unlike our study, the syntactic structures in their long sentence condition employed subject relatives and object relatives, as well as canonical and noncanonical word order. Thus, although their more challenging sentences required processing fewer elements, these sentences were more complex syntactically than in our study. Robertson and Joannis concluded that the children with SLI had difficulty that could be attributed to syntactic factors, but capacity demands (those pertaining to working memory in their view) exacerbated the problem. It would be interesting if in future work, the same long sentences used by Robertson and Joannis could appear in two conditions – one in which the adjectives were superfluous, and the other, as in the present study, where each of

the two attributes had to be retained with its respective noun to ensure a correct response. In this case, the role of syntax could be evaluated in the comparison between canonical long sentences and noncanonical long sentences and the role of cognitive capacity could be assessed through a comparison between long sentences with superfluous adjectives and long sentences with contrastive adjectives.

It can be seen, then, that there are other applications that might help determine where problems with grammar stop and limitations in cognitive capacity begin. We believe that the feature of manipulating capacity demands in tasks involving the same syntactic structure, as we have done here, will prove to be a useful principle in the development of any new research on this important question.

Acknowledgments

This research was supported by grants R01 DC009574 and P30 DC005803 from the National Institute on Deafness and Other Communication Disorders, and grant P30 HD002528 from the National Institute on Child Health and Human Development, National Institutes of Health. The authors would like to thank Lisa Wisman Weil, Windi Krok, Johanna Hassink, and Leigh Hardy for their assistance during various phases of this project.

References

- Archibald L, Gathercole S. Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders*. 2006; 41:675–693. [PubMed: 17079222]
- Bavin E, Wilson P, Maruff P, Sleeman F. Spatio-visual memory of children with specific language impairment: Evidence for generalized processing problems. *International Journal of Language and Communication Disorders*. 2005; 40:319–332. [PubMed: 16195191]
- Bishop DVM. Comprehension in developmental language disorders. *Developmental Medicine and Child Neurology*. 1979; 21:225–238.
- Bishop DVM. Comprehension of spoken, written and signed sentences in childhood language disorders. *Journal of Child Psychology and Psychiatry*. 1982; 23:1–20. [PubMed: 6174536]
- Bishop DVM, Adams C. Comprehension problems in children with specific language impairment: Literal and inferential meaning. *Journal of Speech and Hearing Research*. 1992; 35:119–129. [PubMed: 1735960]
- Bishop DVM, Adams C, Rosen S. Resistance to grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. *International Journal of Language and Communication Disorders*. 2006; 41:19–40. [PubMed: 16272001]
- Bishop DVM, Bright P, James C, Bishop S, van der Lely H. Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics*. 2000; 21:159–181.
- Burgemeister, B.; Blum, L.; Lorge, I. *Columbia Mental Maturity Scale*. New York: Harcourt Brace Jovanovich; 1972.
- Dawson, J.; Eyer, J.; Fonkalsrud, J. *Structured Photographic Expressive Language Test – Preschool*. 2. DeKalb, IL: Janelle Publications; 2005.
- Deevy P, Leonard L. The comprehension of Wh-questions with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2004; 47:802–815.
- Ellis Weismer S, Evans J, Hesketh L. An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1249–1260.
- Finneran D, Francis A, Leonard L. Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*. 2009; 52:915–929.
- Graff Estes K, Evans J, Else-Quest N. Differences in nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*. 2007; 50:177–195.

- Greenslade K, Plante E, Vance R. The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test – Preschool: Second Edition. *Language, Speech, and Hearing Services in Schools*. 2009; 40:150–160.
- Hoffman L, Gillam R. Verbal and spatial information processing constraints in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2004; 47:114–125.
- Kail R. A method of studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech and Hearing Research*. 1994; 37:418–421. [PubMed: 8028323]
- Kail R, Salthouse T. Processing speed as a mental capacity. *Acta Psychologica*. 1994; 86:199–225. [PubMed: 7976467]
- Lee, L. Developmental Sentence Scoring. Evanston, IL: Northwestern University Press; 1974.
- Leonard L, Deevy P. Tense and aspect in sentence interpretation by children with specific language impairment. *Journal of Child Language*. 2010; 37:395–418. [PubMed: 19698206]
- Leonard L, Ellis Weismer S, Miller C, Francis D, Tomblin JB, Kail R. Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research*. 2007; 50:408–428.
- Mainela-Arnold E, Evans J. Beyond capacity limitations: Determinants of word-recall performance on verbal working memory span tasks in children with SLI. *Journal of Speech, Language, and Hearing Research*. 2005; 48:897–909.
- Marton K, Schwartz R. Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2003; 46:1138–1153.
- Miller C, Kail R, Leonard L, Tomblin JB. Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2001; 44:416–433.
- Montgomery J. Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research*. 1995; 38:187–199. [PubMed: 7731209]
- Montgomery J. Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2000; 43:293–308.
- Montgomery J, Evans J. Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2009; 52:269–288.
- Norbury CF, Bishop DVM, Briscoe J. Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics*. 2002; 23:247–268.
- Rispens J, Been P. Subject-verb agreement and phonological processing in developmental dyslexia and SLI: A closer look. *International Journal of Language and Communication Disorders*. 2007; 42:293–305. [PubMed: 17514543]
- Robertson E, Joanisse M. Spoken sentence comprehension in children with dyslexia and language impairment: The roles of syntax and working memory. *Applied Psycholinguistics*. 2010; 31:141–165.
- Spaulding T, Plante E, Vance R. Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research*. 2008; 51:16–34.
- van der Lely H. Domain-specific cognitive systems: Insight from grammatical SLI. *Trends in Cognitive Sciences*. 2005; 9:53–59. [PubMed: 15668097]
- van der Lely H, Harris M. Comprehension of reversible sentences in specifically language impaired children. *Journal of Speech and Hearing Disorders*. 1990; 55:101–117. [PubMed: 2299827]
- van der Lely H, Stollwerck L. Binding theory and grammatical specific language impairment in children. *Cognition*. 1997; 62:245–290. [PubMed: 9187060]
- Wiig, E.; Secord, W.; Semel, E. *Clinical Evaluation of Language Fundamentals – Preschool 2*. San Antonio, TX: Psychological Corporation; 2004.
- Windsor J, Hwang M. Testing the generalized slowing hypothesis in specific language impairment. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1205–1218.

Appendix. Sentences Used in Experimental Task

Low-Demand Items

The horse kicks the cow
The cat kisses the monkey
The bear pulls the elephant
The girl tickles the boy
The monkey feeds the tiger
The bunny touches the turtle
The turtle splashes the frog
The bunny chases the cat
The dog washes the pig
The mouse covers the bird

Intermediate-Demand Items

The nice bear pulls the little elephant
The brown monkey feeds the little tiger
The happy turtle splashes the wet frog
The pretty cat kisses the nice monkey
The happy dog washes the little pig
The nice girl tickles the little boy
The happy bunny chases the orange cat
The gentle bunny touches the nice turtle
The nice mouse covers the pretty bird
The little horse kicks the brown cow

High-Demand Items

The yellow dog washes the white pig
The little bunny chases the happy cat
The brown mouse covers the red bird
The white horse kicks the brown cow
The happy turtle splashes the little frog
The happy bear pulls the sad elephant
The brown monkey feeds the orange tiger
The nice bunny touches the sick turtle
The orange cat kisses the brown monkey
The nice girl tickles the happy boy

Figure 1.
The four drawings used for the item *The yellow dog washes the white pig* from the high demand condition.

Figure 2.
The children's mean sentence comprehension scores on low, intermediate, and high demand items. Error bars reflect 95% confidence intervals.

Table 1

Ages and Test Scores of the Three Groups of Children.

	SLI (N = 19)	TD-A (N = 19)	TD-Y (N = 19) ^e
Age in months			
<i>M</i>	56.21	56.74	40.89
<i>SD</i>	5.64	5.72	3.26
Range	48–66	48–68	37–47
CELF-P2 ^a			
<i>M</i>	13.16	17.32	12.68
<i>SD</i>	2.79	2.33	3.25
Range	9–18	11–22	8–18
CMMS SS ^b			
<i>M</i>	105.47	118.47	113.10
<i>SD</i>	8.21	8.90	13.83
Range	87–118	103–140	90–133
CMMS MI ^c			
<i>M</i>	4.89	6.18	4.40
<i>SD</i>	1.02	1.19	0.94
Range	3.0–7.0	5.0–9.5	3.0–5.5
LIPS ^d			
<i>M</i>			117.11
<i>SD</i>			11.12
Range			100–135

^aRaw score on the Sentence Structure subtest of the CELF-P2.^bStandard score on the Columbia Mental Maturity Scale.^cMaturity Index from the Columbia Mental Maturity Scale.^dStandard score on the Leiter International Performance Scale.^eN = 10 for the CMMS and N = 9 for the LIPS.

Table 2

Number of Correct Responses on the Three Sets of Screening Items

Set	SLI	TD-A	TD-Y
Simple S-V-O (10 items)			
<i>M</i>	9.11	9.16	8.42
<i>SD</i>	0.81	1.01	1.12
Adjectives (20 items)			
<i>M</i>	19.32	19.79	19.05
<i>SD</i>	1.06	0.54	1.08
S-V-O with Superfluous Adjectives (10 items)			
<i>M</i>	9.05	9.32	8.68
<i>SD</i>	0.97	0.67	1.16

Table 3

Mean Accuracy Scores and Arcsine Transformed Scores (and Standard Deviations) for the Three Groups of Children on the Low Demand, Intermediate Demand, and High Demand Items.

Condition	SLI	TD-Y	TD-A
Low Demand			
Accuracy Score	7.84 (1.61)	8.21 (1.75)	9.26 (0.73)
Transformed Score	0.9605 (0.3010)	1.0695 (0.3699)	1.2793 (0.2641)
Intermediate Demand			
Accuracy Score	7.42 (1.71)	7.37 (1.89)	9.05 (1.08)
Transformed Score	0.8965 (0.3194)	0.8904 (0.3300)	1.2566 (0.3215)
High Demand			
Accuracy Score	5.95 (1.99)	6.00 (2.54)	9.00 (0.88)
Transformed Score	0.6754 (0.3055)	0.7142 (0.4057)	1.2036 (0.2743)

Table 4

Distribution of Errors for Each Set of Items. Values are Mean Number of Errors and the Percentage of Errors That the Numbers Represent.

Condition	SLI	TD-Y	TD-A
Low Demand			
Reversal	25 (61%)	18 (53%)	6 (43%)
Incorrect Object ^a	11 (27%)	3 (9%)	3 (21%)
Incorrect Subject ^b	5 (12%)	13 (38%)	5 (36%)
Intermediate Demand			
Reversal	27 (55%)	33 (64%)	11 (61%)
Incorrect Object ^a	12 (24%)	8 (19%)	4 (22%)
Incorrect Subject ^b	10 (20%)	9 (17%)	3 (17%)
High Demand			
Reversal	12 (16%)	21 (27%)	3 (16%)
Incorrect Object ^c	31 (40%)	27 (34%)	6 (32%)
Incorrect Subject ^d	34 (44%)	28 (38%)	10 (53%)

^aSelection of a foil depicting an un-named character in object position.

^bSelection of a foil depicting an un-named character in subject position.

^cSelection of a foil depicting an appropriate character with an inappropriate attribute in object position.

^dSelection of a foil depicting an appropriate character with an inappropriate attribute in subject position.