Syntactic Structural Assignment in Brazilian Portuguese-Speaking Children With Specific Language Impairment

Talita Fortunato-Tavares, a,b Claudia R. F. de Andrade, b Debora M. Befi-Lopes, b Arild Hestvik, c Baila Epstein, d Lidiya Tornyova, e and Richard G. Schwartz e

Purpose: In this study, the authors examined the comprehension of sentences with predicates and reflexives that are linked to a nonadjacent noun as a test of the hierarchical ordering deficit (HOD) hypothesis. That hypothesis and more modern versions posit that children with specific language impairment (SLI) have difficulty in establishing nonadjacent (hierarchical) relations among elements of a sentence. The authors also tested whether additional working memory demands in constructions containing reflexives affected the extent to which children with SLI incorrectly structure sentences as indicated by their picture-pointing comprehension responses.

Method: Sixteen Brazilian Portuguese-speaking children (8;4–10;6 [years;months]) with SLI and 16 children with typical language development (TLD) matched for age (±3 months), gender, and socioeconomic status participated in 2 experiments (predicate and reflexive interpretation). In the reflexive experiment, the authors also manipulated working memory demands. Each experiment involved a 4-choice picture selection sentence comprehension task.

Results: Children with SLI were significantly less accurate on all conditions. Both groups made more hierarchical syntactic construction errors in the long working memory condition than in the short working memory condition.

Conclusion: The HOD hypothesis was not confirmed. For both groups, syntactic factors (structural assignment) were more vulnerable than lexical factors (prepositions) to working memory effects in sentence miscomprehension.

Key Words: specific language impairment, syntax, structural assignment, working memory, children

One characteristic of specific language impairment (SLI) is a deficit in the comprehension and production of sentences, particularly those that involve long-distance relationships. Several proposals have focused on the computational grammatical complexity (CGC) hypothesis, which posits deficits in the computational grammatical system that affect hierarchically complex structures in one or more components (syntax, morphology, phonology) of language (Gallon, Harris, & van der Lely, 2007; Marinis & van der Lely, 2007; Marshall & van der Lely, 2006; van der Lely, 2005; van der Lely & Stollwerk, 1997).

The hierarchical ordering deficit (HOD) account (Cromer, 1978), an historical predecessor of CGC, suggests that children with language impairment do not assign structure to sentences in the same way as their typically developing peers. According to this account, children with language impairment represent sentences as flat instead of hierarchically organized structures, and the absence of hierarchical relations within the sentence structure would account for their overall difficulties with sentence comprehension. The HOD account concerns children with language disorders, but because the discussion and definition of SLI had not yet emerged by the late 1970s, this hypothesis has not been directly studied in children with SLI.
Bishop (1982) examined the HOD hypothesis in children with Landau-Kleffner syndrome who may exhibit some similar language deficits to children with SLI (Billard, Fluss, & Pinton, 2009; Deonna & Roulet-Perez, 2005; Overvliet et al., 2010). In that study, Bishop tested the relation between a predicate and the syntactic position of the nominal to which the predicate applies (attachment). An offline sentence comprehension task was used, in which children heard a sentence and then pointed to a picture in an array of four. Ten sentences with the following structure were devised: The X in/on/under/in front of/behind the Z is Y, where X and Z were nouns and Y was a color term. For each sentence, the corresponding of four pictures were as follows: correct picture (C); sequential error where X and Y are reversed (S); prepositional error, corresponding to a sentence with a different preposition (P); and hierarchical error, corresponding to the type of error where the complement applies to Z rather than X (H). The actual group and individual scores were never presented. Instead, Bishop classified the children as C-, S-, P-, or H-responders according to the most frequent type of response given to the 10 sentences. Children who did not show a preference for selecting one type of picture were classified as mixed. More children with Landau-Kleffner syndrome were classified as having a hierarchical error preference than for any other response. In contrast, children with typical language development (TLD) did not show a particular preference for one type of distractor. Although Bishop concluded that these findings supported the HOD hypothesis as a cause for the deficit in comprehending sentences with hierarchical long-distance dependencies, the small number of stimuli and the unusual data analyses render the conclusion uncertain.

In several studies, researchers have investigated the ability of children with SLI to comprehend and produce sentences with long-distance dependencies. Although there seems to be an agreement that children with SLI do have an overall difficulty with a variety of syntactic structures such as relative clauses (Friedmann & Novogrodsky, 2004, 2007; Hestvik, Schwartz & Tornyova, 2010; Schuele & Tolbert, 2001) and wh-questions (Deevy & Leonard, 2004; Hansson & Nettelbladt, 2006; Marinis & van der Lely, 2007), the overall source of these difficulties is not yet clear. Among the candidate deficits that might explain these difficulties are specific grammatical deficits (e.g., Cromer, 1978; van der Lely, 2005; van der Lely & Stollwerck, 1997) and working memory processing limitations (e.g., Deevy & Leonard, 2004; Hestvik et al., 2010; Marton, Schwartz, Farkas, & Katsnelson, 2006). In the present study, we examined the HOD hypothesis and the effects of working memory demands on syntactic structural assignment of children with SLI and their chronologically matched controls.

**Syntactic Aspects of the HOD Hypothesis**

The HOD hypothesis (Cromer, 1978) is based on the dependency among elements of a syntactic arrangement. When two or more elements co-occur in a syntactic arrangement, some type of dominance exists between or among them. Typically, there is one dominant element, the head, which is the primary determinant of the properties of the arrangement. The other elements are its dependent(s) (Van Valin, 2001). C-command (deriving from constituent command) is a relationship between phrasal nodes (elements of syntactic structure) in grammatical trees. Originally defined by Reinhardt (1976), it corresponds to the idea of siblings and all their descendants in a hierarchical relationship.

To illustrate, when a sentence such as The chicken on the ball is brown is heard, a hierarchical representation (see Figure 1A) of the sentence is incrementally built as each word is processed. However, suppose that children with SLI fail to build the correct syntactic structure during online sentence processing. The same sentence would then be represented as a sequence of units in a flat structure (see Figure 1B), and the absence of hierarchical relations results in an indeterminate attachment. It also leads to an absence of structural differentiation information among the elements of the syntactic tree resulting in comprehension errors with discourse consequences.

In Figure 1A, the subject noun phrase (NP; the chicken) is positioned higher in the tree than the predicate (brown). Therefore, the interpretation of this hierarchically built structure would be that the chicken is on the ball, and

![Diagram of hierarchical structure](image)

**Figure 1.** A: Hierarchical structure representation of the sentence with predicate attachment. The chicken on the ball is brown. B: Incorrect (flat) structure representation of the sentence The chicken on the ball is brown.
the chicken is brown because the subject (the chicken) c-commands the predicate (brown). In Figure 1B, this differentiation is not present and is not expressed by the structure. This lack of hierarchical structure results in ambiguity; the predicate (brown) could be attached either to the nearest NP (the ball) or to the furthest NP (the chicken). In case a recency strategy is employed, the predicate (brown) is applied to the nearest NP (the ball). In the latter, the interpretation of this flat structure would be that the chicken is on the ball, and the ball is brown; in other words, the chicken is not higher than brown in this incorrect syntactic structure.

The supposed incorrect structure assignment proposed by Cromer (1978) should have the same effects on predication relationships and on reflexive interpretation. Reflexive interpretation, similar to predication, relies on structure assignment because a reflexive pronoun must agree with the antecedent that c-commands it in person, gender, and number. C-command essentially means that the antecedent must be asymmetrically higher in the tree representation than the reflexive. For example, in the following syntactic representation (see Figure 2A), the reflexive pronoun can refer only to the grandma, not the mom. The reason for this is that the NP the grandma is the head of the subject. Only the grandma c-commands the reflexive; thus, only this NP can provide reference for the reflexive. If SLI children do not correctly represent these structural relations (as in Figure 2B), their rules for reflexive interpretation would not apply correctly, and they might allow the mom to be the antecedent of the reflexive, yielding the following interpretation: The grandma is behind the mom, and the mom is looking at herself.

**Syntactic Structural Assignment and Reflexives**

The authors of several studies have examined whether deficits in the syntactic structural assignment affect reflexive–antecedent relations in the SLI population using the Syntactic Test of Pronominal Reference (STOP; e.g., van der Lely, 1998; van der Lely & Stollwerck, 1997). The STOP analyzes syntactic assignment of pronouns and reflexives in the following relations: name–pronoun, quantifier–pronoun, name–reflexive, and quantifier–reflexive. The test consists of 108 sentences: 48 experimental sentences and 60 fillers. Each sentence is presented twice. In one presentation, the picture matched the target sentence; in the other presentation, the picture did not match the target sentence.

In a study using the STOP (van der Lely & Stollwerck, 1997), 12 children ages 9.3–12.10 (years; months) with grammatical SLI were near ceiling on the name–reflexive conditions in which semantic–conceptual lexical knowledge was sufficient to identify the correct antecedent. In contrast, children with SLI had significantly lower accuracy scores than did their language-matched controls (ages 5.5 to 8.9) in mismatch conditions with lexical information that was insufficient to determine the antecedent, and the children had to rely only on syntax. In a subsequent study, van der Lely (1998) reported similar findings on a case study of a boy with SLI, but much less details of findings were provided. On these two studies, the authors concluded that children with SLI have sufficient knowledge of the semantic–conceptual properties of reflexives and theta-role assignment, but they do not have the syntactic knowledge characterized by the binding theory.

However, some facts about these studies should be noted. The test included syntactic constructions with subordinate clauses and question structures, which might have added greater complexity to the reflexives. Furthermore, we used a limited number of different sentences per condition (only six, as each sentence was presented twice). Children with SLI were up to 5.5 SDs below the mean on standardized tests—an atypical study group—which also resulted in a very young group of language-matched controls. Moreover, van der Lely and Stollwerck (1997) observed no between-group differences on the match condition in which children had to rely only on syntax. Processing limitations such as working memory deficits may also play a significant role on the construction of hierarchical syntactic structures, and none of these studies had taken that into account.
The role of structured syntactic assignment and the influence of working memory on sentence comprehension of children with SLI still remain unclear.

**Working Memory and Sentence Comprehension**

Although domain-specific perspectives of SLI, such as the CGC account, predict a pervasive deficit in grammatical components determined by structural complexity, they reject the influence of working memory. Working memory is critical for processing language because the building of syntactic and discourse structures requires relating linguistic units across a number of intervening words and syllables in a lengthy time span (Martin & McElree, 2009; Marton et al., 2006; McElree, Foraker, & Dyer, 2003).

Most investigations of the relation between working memory and sentence comprehension have analyzed the correlation between independent working memory measures (e.g., nonword and real-word repetition, repetition of word lists, etc.) and performance on a sentence comprehension task. For example, Just and Carpenter (1992) used a sentence list recall task to classify subjects as having high or low sentence memory spans and then found that their memory span was related to their performance on an independent comprehension task. In contrast, Waters and Caplan (1996) found that sentence list recall span was not a major determinant in the processing of garden-path sentences. This latter finding led to a proposal of multiple working memory capacities that subserve language processing and to the suggestion that the sentence span task does not assess the working memory used for language comprehension in adults.

In several studies, Montgomery and colleagues examined the influence of working memory on sentence comprehension by children with SLI. An early study (Montgomery, 1995) provided evidence of a relation between phonological working memory deficits and sentence comprehension difficulties for some children with SLI. More recent studies (Montgomery, 2000, 2004; Montgomery & Evans, 2009) provided evidence that sentence comprehension is associated with working memory and claimed that deficits in comprehending complex sentences cannot be explained by a syntax-specific representational deficit. These conclusions were also based primarily on the correlation of independent working memory measures and sentence comprehension performance. The working memory tasks that are typically employed reflect the temporary storage of verbal material that plays only a secondary role in higher-level language comprehension. It is not surprising that performance on these tasks is correlated, but to better understand the role of working memory in sentence comprehension, we must use tasks that directly manipulate working memory demands in sentences in order to be more informative.

The few studies that have used such a direct manipulation of working memory load in sentences have posited working memory as one of the underlying causes of sentence comprehension problems in children with SLI. In a study of wh-question comprehension with direct manipulation of sentence working memory demands (Deevy & Leonard, 2004), children with SLI performed similarly to children with TLD on short questions but showed poorer performance on long subject questions than on long subject questions. The combination of a more difficult structure with additional length posed more problems for children with SLI. The authors then concluded that working memory places demands on linguistic processing and thus plays an important role in the difficulties experienced by children with SLI. However, the actual difference between conditions was quite small (one to two items) despite the significant statistical findings.

Marton and colleagues (2006) reported that the increasing in the number of words in a sentence, without an increase in grammatical complexity, did not influence performance accuracy to the same extent as did the increase in morphological complexity. In that study, children correctly recalled more words and answered more questions following sentences with simple morphological structures when compared with sentences that contained complex morphology. These results and previous findings (Marton & Schwartz, 2003) suggest an influence of linguistic processing on working memory demands, with a larger impact of linguistic complexity than sentence length on working memory performance accuracy.

**Overview**

In the present study, we examined whether children with SLI have a deficit in their structural assignment. According to the HOD hypothesis (Cromer, 1978), children with language impairment construct a flat rather than a hierarchical representation of sentences leading to comprehension errors. If the HOD is correct, children with SLI will exhibit lower accuracy and select more errors reflecting a nonhierarchical construction than will children with TLD.

We conducted two experiments to examine two types of structural assignment in comprehension: predicates and reflexives. The first experiment focused on predicate–NP relations as in Bishop (1982). The second focused on reflexives and their antecedent nouns. If the HOD hypothesis is true, it should be valid for both predicates and reflexives.

If the deficit is domain specific as in the CGC and the HOD accounts, working memory effects on sentence
comprehension would not occur. We examined this in the second experiment by manipulating sentence length without adding structural complexity. By examining error patterns, we also expected to be able to tease apart syntactic and lexical factors from working memory effects.

**Cross-Linguistic Issues**

Although the hierarchical structure of the phrases is assumed to be universal, the linear ordering between the head and its complements, as well as between the intermediate level and the specific position, is language specific. The latter is considered a parameter of Universal Grammar, giving rise to various word orders across languages. Brazilian Portuguese is a head-initial language (as are English, French, Hebrew, and many others) in which the head (e.g., the verb *ate*) precedes the complement (e.g., the NP *an apple* in the sentence *John ate an apple*, from Botwinik-Rotem & Friedmann, 2009). There is a high attachment preference in Brazilian Portuguese (Miyamoto, 1998) for both structures under investigation in this study (predicate attachment and reflexive assignment). Thus, it is reasonable to generalize findings from the present study to other head-initial languages and languages with high attachment preference for predicates and reflexives.

Brazilian Portuguese primarily uses reflexive pronouns as clitics attached to the verb as in other Romance languages. Unlike English reflexive pronouns, reflexive clitics in Brazilian Portuguese are unstressed. They have person and number markers (but no gender markers) and cannot be used in isolation. The basic pattern for clitic placement in modern Brazilian Portuguese is preverbal. The Brazilian Portuguese third-person reflexive clitic used in the current study is *se*, as shown in the following example:

A avó atrás da mãe *se* olhou. [The grandma behind the mom looked at herself.]

Bedore and Leonard (2001) observed that Spanish-speaking children with SLI performed more poorly than both age- and mean length of utterance (MLU)–matched controls in marking gender and number of clitics. As no studies have examined clitic production by Brazilian Portuguese-speaking children with SLI, we aimed to exclude the possible bias of clitic morpheme agreement by presenting only one clitic form: Both possible antecedents of clitics in the present study were third person singular (see the example in Figure 3).

Children with TLD have problems in acquiring certain clitic forms in some languages. Frequent clitic omissions in obligatory contexts and late onset of clitics in the object position have been observed for languages such as Brazilian (Lopes, 2003) and European (Silva, 2008) Portuguese, Italian (Guasti, 1993; Schaeffer, 1997), French (Jakubowicz & Rigaut, 2000; Van der Velde, Jakubowicz, & Rigaut, 2002), and Catalan (Wexler, Gavarró & Torrens, 2004). However, determinants with the same morphological form of third person clitics are not omitted. For these languages, clitic omission does not occur because of perceptual or prosodic production factors but because of their syntactic roles.

Although reflexive clitics are less problematic for children than other clitics (Costa & Lobo, 2007; Jakubowicz & Rigaut, 2000; Silva, 2008), the presence of clitics, and not full pronouns such as reflexives in English, for example, still may pose additional difficulties for children other than the pure reflexive assignment examined here. If the presence of clitics posed an additional challenge, we should observe lower accuracy for the reflexive than for the predicate experiment with similar (short) working memory demands.

**General Method**

**Participants**

Thirty-two Brazilian Portuguese-speaking children (19 boys, 13 girls) participated in the two experiments. Children were between the ages of 8;4 and 10;6 and were divided into two aged-matched, equally sized groups (TLD and SLI). All children were paired by age within a 3-month maximum interval by gender (with the exception of one pair) and by socioeconomic status (SES).
according to the Brazilian Economic Classification Criterion questionnaire (Crite\'rio de Classifica\'c\'ao Econ\'omica Brasil [CCEB]; Associa\'c\'ao Brasileira de Empresas de Pesquisa [ABEP], 2008). Children were also matched by educational level: The same number of children were enrolled in third grade (four children), fourth grade (nine children), and fifth grade (three children) in each group. All children came from homes in which Brazilian Portuguese was the only language spoken.

**SLI group.** These 16 children (10 boys and six girls) were recruited through the Investigation on Specific Language Impairment Laboratory at the Department of Physical Therapy, Communication Sciences and Disorders and Occupational Therapy of the Medical School of Universidade de S\’ao Paulo (Faculdade de Medicina da Universidade de S\’ao Paulo), Brazil. The children with SLI were between ages 8:4 and 10:6 (M = 9.4; SD = 0.9). All of the families were identified as C2 (lower middle class) on the CCEB questionnaire. These children had no history of neurological impairments, no evidence of oral motor disabilities, and no social or emotional difficulties. They all had been diagnosed by a group of speech-language pathologists as having SLI. There is no comprehensive standardized language test in Brazilian Portuguese. The most comprehensive language test available is the ABFW Child Language Test (Andrade, Befi-Lopes, Fernandes, & Wertzner, 2004), which evaluates vocabulary, phonology, fluency, and pragmatics. The ABFW is the most widely used language test in Brazil; it is also used in Portugal. All children from the SLI group had ABFW scores that were at least 1.25 SDs below the mean on vocabulary and fluency measures within normal limits regarding number of stuttering-like disfluencies and percentage of stuttered syllables. Children who had only liquid simplification, cluster simplification, or distortions were included in the study. Other phonological processes (e.g., stopping, deaffrication) were considered as exclusion criteria. All children exhibited a predominance of use of verbal as compared to gesture and nonverbal communication means (see Table 1 for more detailed information on test scores). The mean length of utterance in morphemes (MLU<sub>m</sub>) in spontaneous speech samples of at least 150 utterances for all children was below at least 1 SD of reference values available for Brazilian Portuguese (Araujo, 2007). The group mean for MLU<sub>m</sub> was 4.84 (SD = 0.48).

Additional criteria for children with SLI selection included normal nonverbal IQ performance (score above 85) as measured by the Test of Nonverbal Intelligence—Third Edition (Brown, Sherbenou, & Johnsen, 1997) and normal hearing as measured by hearing screening at 25 dB HL for the frequencies of 500 Hz, 1000 Hz, and 2000 Hz. All children in the SLI group had persistent histories of language impairment after more than 2 years of speech and language intervention and were receiving speech-language services at the time of testing.

**TLD group.** Children in this group were recruited from public schools of S\’ao Paulo. The 16 children (nine boys and seven girls) were between ages 8:5 and 10:6 (M = 9.4; SD = 0.8). According to the SES questionnaire, all families of children of this group were classified as C2 on the CCEB (ABEP, 2008) with the exception of one family, which was classified as C1 (middle class). These children had scores on the four sections of the ABFW Child Language Test (Andrade et al., 2004) within normal limits. The MLU was not computed in this group due to limits on testing imposed by most of

<p>| Table 1. Mean (and SD) of age, nonverbal IQ, and language test performance of participants. |
|-----------------------------------------------|---------------|---------------|---------------|</p>
<table>
<thead>
<tr>
<th><strong>Variable</strong></th>
<th><strong>SLI</strong></th>
<th><strong>TLD</strong></th>
<th><strong>p</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>9.4 (0.9)</td>
<td>9.4 (0.8)</td>
<td>.293</td>
</tr>
<tr>
<td>TONI-3</td>
<td>97 (5.6)</td>
<td>101 (5.5)</td>
<td>.079</td>
</tr>
<tr>
<td>ABFW Child Language Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary raw score</td>
<td>52 (3.6)</td>
<td>77 (3.8)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Fluency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stuttering-like disfluencies</td>
<td>1.4 (1.0)</td>
<td>1.3 (0.9)</td>
<td>.580</td>
</tr>
<tr>
<td>Percentage of stuttered syllables</td>
<td>0.7 (0.5)</td>
<td>0.7 (0.4)</td>
<td>.580</td>
</tr>
<tr>
<td>Pragmatics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of communicative acts per minute</td>
<td>8.8 (0.9)</td>
<td>10.6 (1.1)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Percentage of verbal communicative acts</td>
<td>67.5 (3.8)</td>
<td>74.5 (6.6)</td>
<td>&lt;.001*</td>
</tr>
</tbody>
</table>

Note.  SLI = specific language impairment; TLD = typical language development; TONI-3 = Test of Nonverbal Intelligence—Third Edition.

*Statistically significant p values for t tests.
the schools from which children were recruited. Their teachers were questioned about academic performance related to syntax and morphosyntax, and none of the teachers reported difficulties in these areas for any participant. All children from this group had no history of language impairment, normal nonverbal IQ performance (score above 85) as measured by the Test of Nonverbal Intelligence—Third Edition (Brown et al., 1997), and normal hearing as measured by a hearing screening.

**Experiment 1**

In this experiment, we tested the comprehension of predicate–NP relations (attachment) by children with SLI and their matched peers.

**Stimuli**

There were 26 trials in this experiment. Each trial consisted of one context sentence, one target sentence, and an array of four pictures. Context sentences had the following structure: *Here is a(an) X and a(a) Z [Aqui está a(o) X e a(o) Z]*, where X and Z were the nouns of the target sentence (e.g., *Here is a chicken and a ball — Aqui está a galinha e a bola*). Target sentences of this experiment had the following structure: *The X in/on/under/in front of/behind the Z is Y [O X na(o)/acima/abaixo/na frente/atrás de Z é Y]*, where X and Z were nouns and Y was a color term (e.g., *The chicken on the ball is brown — A galinha acima da bola é marrom*). Each visual stimulus included four pictures presented on a computer screen: correct picture (correct attachment or antecedent and correct spatial relation), hierarchical error picture (incorrect attachment or antecedent testing for a hierarchical structural error), preposition change error picture (correct attachment or antecedent with a lexical error on the prepositional relation), and reverse error picture (incorrect attachment or antecedent and spatial relations testing for a complete reversal of relations). Figure 4 illustrates the visual (picture) stimuli for the above trial.

**Procedure**

The stimuli were presented via E-Prime Experimental Control software on a laptop computer. The auditory (context and target sentences) stimuli were digitally recorded by a female Brazilian Portuguese native speaker via Praat software (Boersma & Weenink, 2006) and were presented through the computer speakers at a comfortable audible level.

The trials were randomly presented to avoid length, order, or familiarization effects. The experimenter (first author) verbally provided the following instructions, which were also presented in written form on the computer screen before the beginning of the experiment: *First you are going to listen to a sentence introducing the characters. Four pictures with those characters will appear on the computer, and you will listen to another sentence. You will have to point to the picture that shows what the sentence is about. Five practice trials were conducted. Although we planned to present the practice trials again if a child did not seem to understand the task, all of the children successfully completed the practice trials.

For each trial, the child was presented with a context sentence followed by the visual (four-picture array) stimulus and the target sentence. For example, on a single trial, they heard a context sentence (e.g., *Here is a circle and a star — Aqui está um círculo e uma estrela*). Each context sentence had a maximum duration of 5,500 ms. After an interstimulus interval of 1,000 ms, the target sentence and an array of four pictures were presented simultaneously. The target sentences had the maximum duration of 5,000 ms (e.g., *The circle in the star is blue — O círculo na estrela é azul*). The picture position in the four quadrants of the computer screen was randomly selected for each trial. The four pictures remained on the computer screen until a response was detected. The child then had to select (by pointing) the picture that corresponded to the sentence (the experimenter pushed the corresponding response button). The responses were classified and analyzed according to picture selection (correct, hierarchical error, preposition error, or reversed error).

**Data Analysis**

A widely used approach to analyze data from picture selection tests is to consider the responses as continuous
responses, especially for the TLD group. An appropriate, although less commonly used, statistical model for this type of data is the Dirichlet–Multinomial model (Molenberghs & Verbeke, 2005; Paulino & Singer, 2006). In our analysis, we fitted the Dirichlet–Multinomial model via maximum likelihood methods and compared the expected response frequencies between and within groups via Wald statistics.

Results

Figure 5 illustrates the responses (in percentage) according to picture type selection (correct, hierarchical, preposition change, and reversed) for both groups (TLD and SLI). Children with SLI had a notably lower percentage of correct responses and selected more of each error-type picture when compared to their TLD chronologically matched controls. We employed the Dirichlet–Multinomial model to estimate the probabilities of each response category (correct, hierarchical, preposition change, reversed) for each group (see Table 2). The analysis via Wald statistics revealed an overall effect for group, $\chi^2(3, N = 32) = 117.31, p < .001$, indicating that the response distributions for children with TLD and children with SLI were, in general, different on the comprehension of syntactic constructions involving a predicate attachment.

To further investigate this effect, we used between-group comparisons to examine each response type. Children with TLD were significantly more accurate than children with SLI, $\chi^2(1, N = 32) = 19.41, p < .001$. Moreover, the differences on selection of incorrect pictures presented by the two groups were statistically significant for each of the three types of error: hierarchical, $\chi^2(1, N = 32) = 7.85, p = .005$; preposition change, $\chi^2(1, N = 32) = 7.28, p = .007$; and reversed, $\chi^2(1, N = 32) = 83.82, p < .001$. Thus, the overall lower accuracy of the children with SLI compared with their TLD peers was distributed across the error types.

Table 2. Dirichlet–Multinomial model estimated mean percentage of picture selection for both groups (SU and TLD) on Experiment 1 (predicates) and an Experiment 2 (reflexives) by working memory condition.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Group</th>
<th>Working memory condition</th>
<th>Correct</th>
<th>Hierarchical</th>
<th>Preposition change</th>
<th>Reversed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>M (SE)</td>
<td>M (SE)</td>
<td>M (SE)</td>
<td>M (SE)</td>
</tr>
<tr>
<td>Predicates</td>
<td>TLD</td>
<td>Short</td>
<td>93.0 (3.0)</td>
<td>3.0 (2.0)</td>
<td>2.0 (1.0)</td>
<td>2.0 (1.0)</td>
</tr>
<tr>
<td></td>
<td>SU</td>
<td>Short</td>
<td>59.0 (7.0)</td>
<td>16.0 (5.0)</td>
<td>14.0 (4.0)</td>
<td>11.0 (1.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Short</td>
<td>83.0 (3.0)</td>
<td>4.0 (1.0)</td>
<td>6.0 (2.0)</td>
<td>7.0 (1.0)</td>
</tr>
<tr>
<td></td>
<td>TLD</td>
<td>Long</td>
<td>57.0 (5.0)</td>
<td>10.0 (6.0)</td>
<td>6.0 (2.0)</td>
<td>26.0 (2.0)</td>
</tr>
<tr>
<td></td>
<td>SU</td>
<td>Short</td>
<td>60.0 (5.0)</td>
<td>10.0 (3.0)</td>
<td>15.0 (4.0)</td>
<td>15.0 (1.0)</td>
</tr>
<tr>
<td>Reflexives</td>
<td></td>
<td>Long</td>
<td>33.0 (4.0)</td>
<td>20.0 (4.0)</td>
<td>13.0 (3.0)</td>
<td>34.0 (2.0)</td>
</tr>
</tbody>
</table>
For the within-group error analysis, there were no significant differences in the selection of the three error response types for any of the groups: SLI, $\chi^2(2, N = 16) = 2.08, p = .35$; TLD, $\chi^2(2, N = 16) = 0.33, p = .84$. This indicates that neither children with SLI nor children with TLD exhibited a dominant error response. Although children with SLI made more errors than did children with TLD, their error distribution is no different than that of children with TLD.

**Discussion**

In this experiment, we examined the structural assignment of predicates of children with SLI and their matched controls. Children with SLI performed more poorly than their TLD peers, indicating an overall deficit on the comprehension of sentences with nonadjacent predicate–NP relations in this group. However, a preference for the hierarchical error was not observed. Factors other than a HOD also influenced accuracy of these children. Thus, taken together, these findings do not support the HOD account (Cromer, 1978) as an explanation for the sentence comprehension deficit in SLI.

**Experiment 2**

In this experiment, we tested the comprehension of reflexives and their antecedent nouns by children with SLI and their chronologically matched peers. We also examined whether working memory demands affected the extent to which children with SLI incorrectly comprehended sentences.

**Stimuli**

There were 28 trials for the short working memory condition and 28 trials for the long working memory condition. Each trial consisted of one context sentence, one target sentence, and an array of four pictures.

Context sentences had the same structure as those in Experiment 1 (e.g., *Here is a grandma and a mom — Aqui está uma avó e uma mãe*). Target sentences had the following structure: *The X in/on/under/in front of/behind the Z (modifier) is Y* (O X na(o)/acima/abaixo/na frente/atrás de Z (modific.) está Y). In this experiment, X and Z were nouns and Y was verbal phrase with a reflexive pronoun. In addition, each target sentence was presented in a short and a long version, produced by adding a modifier phrase between the subject and the reflexive without increasing structural complexity (e.g., *The grandma behind the mom with the brown skirt is looking at herself — A avó atrás da mãe de saia marrom está se olhando*).

As in Experiment 1, the four visual stimuli were designed to address the configurations of the target sentence (correct picture), the hierarchical error, the preposition change error, and the reverse error. Figure 3 illustrates the picture stimuli of the trial of Experiment 2 exemplified above.

**Procedure**

The procedure of Experiment 2 was identical to that of Experiment 1. Trials of Experiments 1 and 2 were randomly mixed together in a single session to avoid effects of length, order, or familiarization. The duration of the study (Experiments 1 and 2) was approximately 40 min. The session was divided into three blocks. Breaks were provided between the blocks.

**Results**

Ms (in percentage) of responses according to picture type selection (correct, hierarchical, preposition change, and reversed) for both groups (TLD and SLI) on the short working memory condition are displayed in Figure 6. Figure 7 illustrates the results for the long working memory condition. We used the Dirichlet–Multinomial model to estimate the probabilities of each response type (see Table 2) and the Wald statistic to examine between- and within-group differences.

![Figure 6. M percentage of responses according to picture selection for both groups (TLD and SLI) on Experiment 2 (reflexives) on the short working memory condition. Error bars denote 95% CI.](chart.png)
Figure 7. A percentage of responses according to picture selection for both groups (TLD and SLI) on Experiment 2 (reflexives) on the long working memory condition. Error bars denote 95% CI.

Short working memory condition. The Dirichlet-Multinomial model revealed, via Wald statistics, an overall effect for group, \(\chi^2(3, N = 32) = 44.74, p < .001\), indicating that response distributions for TLD and SLI children were different for the comprehension of constructions involving reflexive–antecedent assignment with low working memory demands.

To further investigate this effect, we conducted between-group comparisons for each of the four response types. Children with TLD were significantly more accurate than children with SLI, \(\chi^2(1, N = 32) = 14.27, p < .001\). Moreover, the group differences in selection of incorrect pictures were statistically significant for each of the three types of error: hierarchical, \(\chi^2(1, N = 32) = 4.67, p = .032\); preposition change, \(\chi^2(1, N = 32) = 5.92, p = .015\); and reversed, \(\chi^2(1, N = 32) = 38.91, p < .001\).

In the within-group error analysis, there were no significant differences among the three error types for the SLI group, \(\chi^2(2, N = 16) = 4.24, p = .12\), indicating that the children with SLI did not exhibit a dominant error response. In contrast, the distribution of errors on the TLD group was not homogenous, \(\chi^2(2, N = 16) = 13.68, p < .001\); children with TLD exhibited differences according to the categories of errors. These findings indicate a different pattern of error response distributions for children with SLI and children with TLD for the comprehension of reflexive constructions with low working memory demands.

In additional analyses, we investigated whether children with SLI and children with TLD had a preference for errors involving an incorrect syntactic assignment. The preposition change picture represents choices in which the child has constructed a hierarchical structure for the sentence (only making an error in the prepositional relation of the pictured items). The remaining two error choices, hierarchical and reversed, represent incorrect syntactic assignment. Therefore, on the next set of analyses, we combined the responses in order to obtain two categories: correct structural assignment (correct response and preposition error response) and incorrect structural assignment (hierarchical and reversed error responses).

When we analyzed the correct (correct and preposition error pictures) and incorrect (hierarchical and reverse error pictures) structural assignment by combining the responses, the model revealed within-group differences for each of the groups: TLD, \(\chi^2(1, N = 16) = 471.43, p < .001\); SLI, \(\chi^2(1, N = 16) = 62.93, p < .001\). In the comprehension of reflexive constructions with low working memory demands, children exhibit a preference for responses with correct structural assignment regardless of language status.

Long working memory condition. The Wald statistical analysis of the Dirichlet–Multinomial model revealed an overall effect for group, \(\chi^2(3, N = 32) = 13.00, p = .005\), in the comprehension of sentences containing reflexive–antecedent relations with increased working memory demands. This finding indicates that responses of TLD and SLI children were, in general, different on this condition.

To further investigate this effect, we conducted between-group comparisons for each of the response types. Children with TLD were significantly more accurate than children with SLI, \(\chi^2(1, N = 32) = 12.45, p < .001\). Moreover, the groups differed in their selection of incorrect pictures for each of the three types of error: hierarchical, \(\chi^2(1, N = 32) = 4.33, p = .040\); preposition change, \(\chi^2(1, N = 32) = 3.84, p = .050\); and reversed, \(\chi^2(1, N = 32) = 7.91, p = .005\).

The within-group error analysis revealed significant differences for the SLI group, \(\chi^2(2, N = 16) = 57.48, p < .001\), and for the TLD group, \(\chi^2(2, N = 16) = 95.03, p < .001\). Both groups exhibited a heterogeneous error distribution for the comprehension of reflexive constructions with increased working memory demands.

When we analyzed the correct (correct and preposition error pictures) and incorrect (hierarchical and reverse error pictures) syntactic structural assignment by combining the responses, the model revealed within-group differences for the TLD group, \(\chi^2(1, N = 16) = 9.45, p = .002\). However, children with SLI did not exhibit such differences, \(\chi^2(1, N = 16) = 1.09, p = .29\). This finding
indicates that when working memory demands were higher, children with SLI exhibited no difference between correct and incorrect syntactic assignment on the comprehension of sentences involving a reflexive-antecedent relation.

**Working memory effect.** The Wald statistics applied to the statistical model indicated significant differences on accuracy between the two working memory conditions for the TLD group, $\chi^2(1, N = 16) = 18.48, p < .001$, and for the SLI group, $\chi^2(1, N = 16) = 14.88, p < .001$. Both groups exhibited a decrease in accuracy with an increase in working memory demands.

A set of analyses examined whether children were more likely to make syntactic (hierarchical and reverse) and lexical (preposition) errors in the long working memory condition. In these analyses, we compared the selection of each error type across the two conditions of working memory. Children with SLI exhibited increased selection of hierarchical, $\chi^2(1, N = 16) = 5.04, p = .025$, and reversed, $\chi^2(1, N = 16) = 98.80, p < .001$, pictures in the long working memory condition. In contrast, children with SLI exhibited a nonsignificant decrease in the selection of the preposition error picture in the long working memory condition, $\chi^2(1, N = 16) = 0.37, p = .540$. The same pattern was observed for children with TLD: a significant increase in the selection of the hierarchical, $\chi^2(1, N = 16) = 4.31, p = .037$, and reversed, $\chi^2(1, N = 16) = 71.33, p < .001$, types of error, and no change in the selection of the preposition error, $\chi^2(1, N = 16) < 0.001, p = .99$. The selection of the preposition change picture was the only type of error not responsible for the decreased accuracy noted for both groups on the long working memory condition.

**Discussion**

In this experiment, we examined the structural assignment for sentences with reflexives and the effect of a greater working memory demand on accuracy and error types. As expected, the children with SLI had lower accuracy than their typically developing peers in both working memory conditions. The poorer performance of the children with SLI in assigning the correct antecedent to reflexives are in agreement with findings reported by van der Lely (1998) and van der Lely and Stollwerck (1997). In those earlier studies, the authors reported that children with SLI exhibited deficits on the assignment of the correct antecedent name to reflexive anaphora when syntactic knowledge was required. However, the following factors make comparisons difficult: the smaller number and the increased complexity of stimuli on the STOP, the use of elicitations rather than full pronouns in the current experiment, the differences in subjects (the children in the previous study have very severely impaired children with SLI and very young language-matched controls), and the different tasks employed.

Syntactic deficits were not the only factors that affected comprehension of hierarchically complex structures. The CGC hypothesis (van der Lely, 2005) posits that only deficits in the computational grammatical system affect hierarchically complex structures and rejects the notion that working memory plays a role in comprehension of such structures. The increase in working memory demands in the present experiment significantly affected the comprehension accuracy for both groups. Working memory demands interfered with hierarchical structure assignment regardless of language status.

Although our work provides converging evidence that children with SLI have working memory limitations affecting their comprehension (Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990; Mainela-Arnold & Evans, 2005; Marton & Schwartz, 2003; Montgomery, 2004; Schwartz, 2009), the two groups exhibited a similar decrease in accuracy (approximately 26%) and similar error patterns with increased working memory demands. Both groups made more syntactic errors (i.e., assignment of incorrect antecedent) in the long than in the short working memory condition. Both the hierarchical and the reversed error type involved structural deviations either in the absence of hierarchical relations or reversed syntactic tree components. These errors increased with a more demanding working memory load. In contrast, preposition errors only involved a lexical error and were selected less frequently in response to the longer sentences. The fact that the lexical error choices decreased whereas the syntactic ones increased in the long working memory condition for both groups indicates that working memory has differential effects on language comprehension regardless of language status.

**General Discussion**

We conducted two experiments designed to examine the ability of children with SLI to establish long-distance relationships between predicates and nouns as well as between reflexives and their antecedents. According to the HOD hypothesis (Cromer, 1978), children with language impairment have a deficit in hierarchical sentence structuring that causes comprehension errors. We compared the comprehension performance and the influence of different working memory demands on two syntactic constructions that require hierarchy knowledge of their elements for a correct comprehension between a group of children with SLI and an age-matched group of children with TLD in order to examine this proposal.
The results of Experiment 1 are not consistent with findings reported by Bishop (1982). These between-study inconsistencies may reflect methodological differences (e.g., fewer trials and no statistical analyses). Furthermore, the participants in Bishop’s study were children with Landau-Kleffner syndrome, and although some similarities in language profiles with children with SLI are expected (Billard et al., 2009; Deonna & Roulet-Perez, 2005; Overvliet et al., 2010), there do not appear to be any studies that have examined syntactic aspects of comprehension in these children. Although we did observe a syntactic structural assignment influence on sentence miscomprehension of children with SLI, the selection of the hierarchical picture was not the dominant error response as observed by Bishop. Instead, both groups of children exhibited no apparent preference among the three error types (hierarchical, preposition change, or reversed) on Experiment 1. This contradicts the HO hypothesis that the most common cause of miscomprehension of children with language impairment is a deficit in structuring the syntactic hierarchical representation of the sentence, resulting in incorrect attachment or antecedent selection.

Although children with SLI made more errors than their TLD matched controls, they still got the majority of trials correct in Experiment 1 and in the short condition in Experiment 2. This suggests that other factors (e.g., issues in processing) may underlie sentence comprehension deficits in children with SLI. Previous investigations have shown that children with SLI are influenced by semantic and pragmatic factors when these factors are related to syntactic constraints (Precious & Conti-Ramsden, 1988; van der Lely & Dewart, 1986). Although plausibility might affect attachment and antecedent–reflexive relations, this factor was controlled in our study (in all situations, both characters were possible antecedents; both the grandma and the mom could be looking at themselves). Moreover, semantic and pragmatic factors may have different effects on attachment and on antecedent–reflexive relations. Predicates and reflexives differ in that the adjective used in the attachment is an open-class word, whereas the reflexive pronoun is a closed-class word and, in this case, a clitic. Although such effects could have influenced the two experiments, this was not observed in the response patterns. Similar accuracy and similar error type distributions were observed in Experiment 1 and in the short working memory condition of Experiment 2.

Other cognitive abilities could also have influenced the performance of children with SLI in the present study. Studies have shown that children with SLI exhibit difficulties with certain cognitive tasks besides having nonverbal IQs near the low end of normal (Hick, Botting & Conti-Ramsden, 2005; Kamhi, Catts, Mauer, Apel, & Gentry, 1988; Marton, 2008). Hick and colleagues (2005) found that children with SLI exhibit a visuospatial deficit, but only when the tasks also involve working memory demands. Our tasks did not involve visuospatial working memory; pictures were presented simultaneously and remained in view until a response was detected.

In Experiment 2, although children with SLI did get a number of trials correct, they were less accurate than their age-matched peers. Although the deficit in syntactic assignment of antecedent nouns to reflexives is consistent with findings from van der Lely (1998) and van der Lely and Stollwerk (1997), some differences between our study and these studies should be noted. In the current study, the reflexives were marked by a clitic pronoun, and there has been evidence for a deficit in clitic acquisition for many Romance languages including Brazilian Portuguese (Loes, 2004). Furthermore, it is noteworthy that clitics are unmarked, and the lack of perceptual salience could also have caused an additional difficulty when compared to full pronouns. Although the clitics in the reflexive constructions could have caused additional problems for children with SLI, this does not seem to have occurred. When working memory demands were more limited, the children with SLI were similarly accurate in the predicate and the reflexive experiments. Van der Lely and colleagues employed a forced-choice judgment task (children saw only one picture and judged whether it matched a sentence they heard), which may have been easier than making a selection from a set of four pictures. However, in a yes/no judgment, guessing may have had a greater effect on the results. Although the three studies examined the assignment of an antecedent noun to a reflexive, three different syntactic constructions were used. Our study involved sentences with an NP with a complex subject followed by a verbal phrase (VP) with a reflexive pronoun (e.g., The grandma behind the mom is looking at herself). Van der Lely and Stollwerk (1997) and van der Lely (1998) used a subordinate structure (e.g., Mougli says Baloo Bear is tickling himself) and a question structure (e.g., Is Mougli tickling himself?). The complex subjects in our study might have been more challenging than the simple subjects in the other two studies. Furthermore, the questions used in the studies of van der Lely and colleagues involved verb movement and included only one possible antecedent, whereas the sentences with subordinate clauses added substantial syntactic complexity that may have confounded the comprehension of reflexives. In our study, the control of the syntactic construction, the analysis of the error types presented by children, and the analysis of working memory demands are important advances.

A different pattern of response was observed in both groups across the working memory demands. With an increased working memory demand, the likelihood of a syntactic representational error also increased.
contrast, with the increased working memory demand, we observed no change in the number of lexical (propositional) errors. Thus, there appears to be a working memory effect on structuring syntactic representations that was not predicted by HOD (Cromer, 1978) or by the CGC (van der Lely, 2005; van der Lely & Stollwerck, 1997) proposals. Clearly, sentence comprehension difficulties are not only a matter of structural assignment but also an effect of working memory demands. Importantly, children with TLD also exhibit working memory effects.

Marton and colleagues (2006) have found that working memory deficits are related to increases in morphosyntactic or syntactic complexity rather than to absolute capacity. However, when we controlled for syntactic complexity (long working memory condition sentences of the current study had no additional structural complexity), working memory still plays an important role in sentence comprehension. This supports the view that the effect of working memory demands on linguistic processing, and not just linguistic complexity, plays an important role in the sentence comprehension difficulties experienced by children with SLI (Deevey & Leonard, 2004). This was also true for children with TLD. Thus, there seems to be a direct effect of working memory demand on sentence comprehension in children rather than simply an association of measures reflecting working memory and sentence comprehension performance as suggested by Montgomery and Evans (2009) and Just and Carpenter (1992).

Waters and Caplan (1996) proposed that at least two different working memory capacities subserve language processing. Although we did not examine this directly, we found evidence of different working memory effects for syntactic (hierarchical structure) and lexical (proposition) factors in sentence comprehension. In our study, the hierarchical relations of the syntactic tree components were affected by the high working memory demand to a greater extent than the proposition use. McElree and colleagues (McElree, 1998, 2000; McElree et al., 2003) have proposed that sentence comprehension is subserved by content-addressable memory structures in which syntactic and semantic constraints provide direct access to relevant representations. Although according to this view both syntactic and semantic constraints seem to be available, there might be different effects on sentence comprehension for each of these constraints. Moreover, if linguistic comprehension is related to the central executive component of Baddeley’s model (1986), as suggested by Marton and colleagues (2006), the model may also need to account for these differential effects across language features.

Although there are mixed findings regarding which working memory measure/component is most closely associated with language comprehension, it is clear from our data that direct manipulation of working memory on sentences does have a general effect on syntactic structural assignment. Furthermore, working memory, as measured by direct manipulation of sentences without adding syntactic structural complexity, has differential effects on comprehension errors (syntactic vs. lexical) regardless of language status.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders (NIDCD) Grant R01DC003885 and Professional Staff Council, City University of New York Grant 62744-00 40, both awarded to the seventh author; a grant from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Coordination for the Improvement of Higher Education Personnel) (CAPES), awarded to the first author; and NIDCD Grant R03DC006175, awarded to the fourth author. We thank the children and families who participated, Fabiola Juste for recording the audio stimuli, and Hia Ditta for drawing the picture stimuli. We are also grateful to Julio Singer for his contribution to data analysis.

References


