JSLHR

Research Article

"Whatdunit?" Sentence Comprehension Abilities of Children With SLI: Sensitivity to Word Order in Canonical and Noncanonical Structures

James W. Montgomery,[a] Ronald B. Gillam,[b] Julia L. Evans,[c] and Alexander V. Sergeev[d]

Purpose: With Aim 1, we compared the comprehension of and sensitivity to canonical and noncanonical word order structures in school-age children with specific language impairment (SLI) and same-age typically developing (TD) children. Aim 2 centered on the developmental improvement of sentence comprehension in the groups. With Aim 3, we compared the comprehension error patterns of the groups.
Method: Using a "Whatdunit" agent selection task, 117 children with SLI and 117 TD children (ages 7:0–11:11, years:months) propensity matched on age, gender, mother's education, and family income pointed to the picture that best represented the agent in semantically implausible canonical structures (subject–verb–object, subject relative) and noncanonical structures (passive, object relative).
Results: The SLI group performed worse than the TD group across sentence types. TD children demonstrated developmental improvement across each sentence type, but children with SLI showed improvement only for canonical sentences. Both groups chose the object noun as agent significantly more often than the noun appearing in a prepositional phrase.
Conclusions: In the absence of semantic–pragmatic cues, comprehension of canonical and noncanonical sentences by children with SLI is limited, with noncanonical sentence comprehension being disproportionately limited. The children's ability to make proper semantic role assignments to the noun arguments in sentences, especially noncanonical, is significantly hindered.

Children with specific language impairment (SLI) exhibit normal-range nonverbal intelligence, hearing sensitivity, articulation, and no neurological impairment or developmental disability. Yet they demonstrate significant language problems for their age. One hallmark deficit is in sentence comprehension (Leonard, Deevy, Fey, & Bredin-Oja, 2013; Montgomery & Evans, 2009; Montgomery, Evans, & Gillam, 2009; Robertson & Joanisse, 2010). Children with SLI are likely to miscomprehend a sentence such as *The cat was chased by the mouse around the barn* as the cat being the chaser (agent) and the mouse as the one being chased (patient). Such comprehension difficulties in school-age children are likely to interfere with oral communication and understanding written texts (Stothard & Hulme, 1992), which often contain a high proportion of complex canonical and noncanonical word order sentences (Scott, 2009; Westby, 1994).

Spoken sentence comprehension involves the ability to rapidly build phrase and clause structures in the construction of a syntactic–semantic representation of sentence meaning. Psycholinguistic studies suggest that individuals initiate comprehension from sentence onset (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978; Marslen-Wilson & Zwitserlood, 1989; Zwitserlood, 1989), building a mental representation of structure and meaning on a word-by-word basis (Borovsky, Elman, & Fernald, 2012; Elman, 1990; Traxler & Tooley, 2007) from the available linguistic and nonlinguistic cues. Word order and noun animacy are robust cues to structure and meaning in English (Bates & MacWhinney, 1987, 1989). In this study, we focused on children's use of word order cues to comprehend the propositionality of canonical and noncanonical structures to

[a]Communication Sciences and Disorders, Ohio University, Athens
[b]Communication Disorders and Deaf Education, Utah State University, Logan
[c]School of Behavioral and Brain Sciences, University of Texas–Dallas, Richardson
[d]Social and Public Health, Ohio University, Athens
Correspondence to James W. Montgomery: montgoj1@ohio.edu

Editor: Sean Redmond
Associate Editor: Megan Dunn Davison
Received January 19, 2017
Revision received March 14, 2017
Accepted April 17, 2017
https://doi.org/10.1044/2017_JSLHR-L-17-0025

Disclosure: The authors have declared that no competing interests existed at the time of publication.

determine who did what to whom in sentences in which animacy cues were controlled.

In English, canonical word order is subject–verb–object (SVO: *The monkey bit the lion*). SVOs pose little difficulty for either school-age TD children or children with SLI (Montgomery & Evans, 2009; van der Lely, 1996) because children are able to process them in a linear manner to determine who did what to whom. As the verb phrase is processed, it directly assigns Noun Phrase 1 (NP1), which grammatically functions as the subject, the thematic role of agent and NP2, which functions as the object, the role of patient. Subject-relative (SR) sentences, those that contain an embedded SR clause (*The monkey [that bit the lion] was brown*), are also canonical in nature. Even though SRs include a relative clause (*that bit the lion*), they also tend to pose little trouble for school-age TD children and children with SLI (Friedmann & Novogrodsky, 2004) because the structure conforms to SVO word order with the head NP (*the monkey*) functioning as both the subject of the main clause (*The monkey ran into the jungle*) and subject of the relative clause (*The monkey bit the lion*).

Reversible verbal *be* passives (*The lion* $_{[i]}$ *was bitten* $_{[ti]}$ *by the monkey*) and embedded object-relative (OR) sentences (*The lion* $_{[i]}$ *that the monkey bit* $_{[ti]}$ *was brown*) are noncanonical in nature, which make them more syntactically complex and difficult to comprehend than canonical structures. Although the surface forms of passives (noun–verb–noun, NVN) and ORs (noun–noun–verb, NNV) are different, in both cases, children must come to realize that (a) NP1 appears in the subject position but it functions as a patient and (b) NP2 occupies the object position but it functions as an agent. What ties these structures together syntactically is that both presumably entail movement and the building of a long-distance syntactic (filler–gap) dependency (for a detailed description of syntactic structure and complexity from a minimalist program perspective, for example, see Chomsky, 1995, and Haegeman, 1994). In both structures, there is movement of the logical object NP (*the lion*) of the verb (*bitten, bit*) to the subject position. The canonical relationship of the verb to its object is maintained by the moved element leaving a trace ([t_i]) behind in its original object position, and the trace shares a coreferential relationship with the moved element ([$_i$]). Both structures also presumably entail reactivation of the moved element to establish the filler–gap dependency in that NP1 (filler, marked as [$_i$]) is reactivated after processing the verb (trace/gap site). In passives, the agent role gets transmitted to NP2 from the passive morpheme (*bitten*), and NP1 receives the patient role. In the case of ORs, once the embedded verb is processed, it assigns the role of agent to NP2 and patient to NP1.

Children with SLI, relative to TD children, have significantly greater trouble understanding reversible *be* passives (Montgomery & Evans, 2009; van der Lely, 1996; van der Lely & Harris, 1990; van der Lely & Stollwerck, 1997) and ORs (Friedmann & Novogrodsky, 2004, 2007; Hestvik, Schwartz, & Tornyova, 2010). But like TD children, children with SLI show better comprehension of passives than ORs (Dick, Wulfeck, Krupa-Kwiatkowski, & Bates, 2004). A prominent linguistic explanation of the sentence comprehension deficits of children with SLI has been offered by van Lely and colleagues (e.g., Marinis & van der Lely, 2007; Marshall, Marinis, & van der Lely, 2007). These investigators proposed a syntax-specific deficit account that describes these children as having trouble computing long-distance syntactic dependencies. The account assumes that the representation and/or mechanisms responsible for building such dependencies are not obligatorily used by children with SLI. In the case of movement-derived structures (passives, ORs), these children treat the movement operation as optional. Doing so results in their inconsistently building the dependency and/or making the proper thematic role assignments to each NP, which, in turn, leads to compromised comprehension. The hypothesis, however, has trouble explaining the finding that children with SLI and TD children show comparable comprehension of nonreversible *be* passives in which the agent noun is animate and the patient noun is inanimate—for example, *The milk was spilled by the girl* (van der Lely, 1996). Such findings provide compelling evidence of the importance of semantic–pragmatic cues facilitating the comprehension of noncanonical structures in children with SLI.

Research focusing on the importance of individual linguistic cues to the sentence comprehension of children with SLI independent of other cues (e.g., noun animacy) is nonexistent. Results from focused research efforts on this issue would lead to substantively new theoretical insights into cue-based driving forces behind these children's comprehension abilities. To this end, this study had three aims. The first was to compare school-age children with SLI and same-age TD peers in their ability to use word order to guide the comprehension of two kinds of canonical structures (SVO, SR) and two kinds of noncanonical structures (passive, OR). The sentences, it is important to note, were absent meaningful semantic–pragmatic cues, thus forcing the children to use only word order cues for comprehension. Children were asked to identify the agent/actor in each sentence, indexing understanding of who did what to whom. The second aim was to determine if children with SLI and TD children show developmental improvement in the comprehension of canonical and noncanonical word order structures. The third was to determine if children with SLI and TD children exhibit different patterns of comprehension errors.

Word Order Use by Children With SLI

Dick et al. (2004) were some of the first investigators to compare word order usage by children with SLI with those of TD peers to guide sentence comprehension. The authors studied children between 5 and 18 years old. An agent selection task (e.g., "whodunit") was used to assess children's understanding of canonical sentences (SVO: *The pig is bumping the goat,* SR: *It's the pig that is bumping the goat*) and noncanonical structures (passive: *The goat is bumped by the pig,* OR: *It's the goat that the pig is bumping*).

Children listened to a sentence and saw two images, one of the agent and the other of the patient. They were instructed to select the agent as quickly as possible. All of the sentences contained animate nouns and were semantically reversible.

Results showed that TD children were sensitive to both the canonical and noncanonical word orders and that sensitivity improved with age. They performed near ceiling on SVOs (approximately 98%) and SRs (approximately 97%). On passives, they achieved 94% accuracy and 86% accuracy on the ORs. Regarding age effects, 5- to 6-year-olds comprehended SVOs and SRs with 90% accuracy with at-ceiling performance achieved by age 9–10 years. Passive comprehension approached asymptote by age 11–12 years. ORs were most difficult. The greatest improvement occurred between ages 9 and 12 years; however, adult-like performance was not reached until age 15–17 years.

The SLI group yielded poorer accuracy of each sentence type than the TD group, especially for the noncanonical structures (SVO: approximately 92%, SR: approximately 91%, Passive: approximately 69%, OR: approximately 57%). Most interestingly, comparing the entire SLI group with the youngest TD children (5- to 7-year-olds), the SLI group showed significantly poorer accuracy for both passives and ORs. Even when comparing the oldest children with SLI (10- to 17-year-olds) with the youngest TD children, the children with SLI were still significantly poorer at OR comprehension. By contrast, these same age groups did not differ on the canonical word order structures. These findings, however, must be interpreted cautiously given the very small sample size of each SLI age group (younger = 16, older = 8).

Noun Animacy Use by Children With SLI

Evans (2002) and Evans and MacWhinney (1999) examined the role of noun animacy (along with word order) in the sentence interpretation abilities of children with SLI and TD children. In the Evans study, children were 6–8 years of age. The children with SLI were divided into those who appeared to be in a transition period in which they moved between using either animacy or word order to interpret a sentence and those who exhibited an emergent use of word order over animacy. The children in the Evans and MacWhinney study were 6–7 years old. The children with SLI were divided into those with an expressive deficit (expressive SLI) and those with mixed expressive–receptive deficits (mixed SLI). In both studies, children heard three different word order structures with either an animate or inanimate agent, such as NVN: *The cat touches the mouse, The tree touches the cat*; NNV: *The goat the horse kisses, The fence the horse kisses*; VNN: *Kisses the bird the dog, Hugs the chair the dog*). Children saw two images (agent image, patient image) and were asked to select the one that was "doing the action." Although the NNV and VNN items were unusual, they permitted the authors to examine the children's reliance on animacy or word order as a cue to agency. The idea was that if word order strategies were beginning to emerge, one would predict NP1 choice to be high regardless of the animacy cues for highly probable sentences, such as NVN, but not for unfamiliar NNV or VNN constructions. If NP1 choice was high only when NP1 was animate in NNV or VNN items, it would suggest the children reverted to an animate-noun-as-agent strategy. If children were primarily relying on a strategy of animate noun as agent, one would predict high NP1 choice when NP1 was animate and low NP1 choice when NP2 was animate regardless of word order.

Evans and MacWhinney (1999) found that the expressive SLI group relied heavily on an NP1 agent comprehension strategy irrespective of animacy whereas the mixed SLI group primarily used animacy to determine agency with performance near chance when animacy cues were absent. The TD children used word order to assign agency regardless of the availability of animacy. In Evans (2002), the children with SLI in transition (nonanimacy cue users) showed sensitivity to word order by selecting NP1 as agent in the NVN and NNV constructions and NP2 as agent for the VNN forms. Those children not in transition (the animacy cue users) always picked the animate noun regardless of the presence of word order cues. The TD children, in contrast, used word order to determine agency, with NP1 selection most often in NVN structures and least often in the NNV structures regardless of animacy cues. Such findings suggest that children with SLI (especially with mixed SLI) are less sensitive to word order cues than same-age TD peers.

Aims of the Present Study

Children with SLI have difficulty using word order to guide sentence comprehension especially for noncanonical structures. These children primarily rely on animacy to determine who did what to whom even through age 8 years (Evans, 2002). Previous studies, however, have not been designed to directly examine these children's word order usage independent of semantic–pragmatic cues. The present study was designed to do just this.

There were three aims. The first was to determine if children with SLI and same-age TD peers differ in their ability to comprehend canonical (SVO, SR) and noncanonical (passive, OR) structures when only word order cues were available. The second was to determine if both children with SLI and TD children demonstrate developmental improvement in comprehending such structures. The third aim was to determine if the groups differed in their comprehension error patterns. To address these aims, 7- to 11-year-old children with and without SLI listened to semantically implausible sentences and selected the agent of the sentence from an array of three images (image of agent, image of patient, image of noun appearing in a sentence-final prepositional phrase) presented at sentence offset. Implausibility was created by controlling the availability of semantic and pragmatic cues (noun animacy, natural affordance between nouns), forcing children to rely on word order cues for comprehension.

We first predicted that the children with SLI would yield poorer comprehension than TD children on both canonical and both noncanonical sentence types (Dick et al., 2004). Second, we expected that, like TD children, children with SLI should show developmental improvement in the comprehension of each sentence type given our relatively large sample sizes. Third, we anticipated that both groups would make predominately syntactically based comprehension errors—that is, assign agency to the object noun as opposed to the noun in the prepositional phrase.

Method

Participants

Participants included children with SLI and TD children who took part in a larger ongoing, multisite project investigating the relationship of cognitive processing and sentence comprehension. The current study centers on the sentence comprehension of children with SLI and same-age TD children. Subsequent reports will detail the cognitive underpinnings of sentence comprehension in these groups. Participants in this study were 234 children, ages 7 through 11 years, 117 with SLI (age_M = 9:5, years:months) and 117 TD children (age_M = 9:5). This age band was chosen because children with SLI demonstrate significantly poorer sentence comprehension than TD peers across these ages, especially for noncanonical structures. Children were recruited from four regions of the United States: Athens, Ohio; Logan, Utah; San Diego, California; and Dallas, Texas. Children were recruited through various school systems, community centers, and university-sponsored summer camps for children.

The degree of exposure to a second language was strictly controlled with English being the primary language spoken by all the children. Similar to procedures used by Bedore et al. (2012), parents were asked to provide a detailed account about their children's language use at home and at school. Bedore et al. found that measures of English semantics and morphosyntax in a large sample of bilingual kindergartners were not affected until children spoke a second language approximately 80 min (20%) each day. Taking a conservative approach, we excluded any child who spoke more than an average of 30 min of another language in the home or at school each day.

All of the children had normal medical history and no neurological impairment or psychological/emotional disturbance on the basis of parent report. They also showed (a) normal-range nonverbal intelligence as indexed by their average score on four nonverbal subtests from the visualization and reasoning battery of the Leiter International Performance Scale–Revised (Roid & Miller, 1997), (b) normal-range hearing sensitivity bilaterally for the frequencies 500 Hz through 4 kHz (American National Standards Institute, 1997), (c) normal-range articulation on the word articulation subtest of the Test of Language Development–Primary: Fourth Edition (Newcomer & Hammill, 2008), and (d) normal or corrected vision.

SLI Group

Consistent with multidimensional systems for defining SLI (e.g., Leonard, 2014; Tager-Flusberg & Cooper, 1999), children were deemed to have significant difficulties in language comprehension or production if they attained a composite language score falling at or below −1 *SD* from the mean on measures of receptive and expressive abilities in the lexical and sentential domains. The four language measures included in the composite were the receptive and expressive portions of the Comprehensive Receptive and Expressive Vocabulary Test—Second Edition (CREVT-2, Wallace & Hammill, 2000) and the concepts and following directions and recalling sentences subtests of the Clinical Evaluation of Language Fundamentals—Fourth Edition (CELF-4, Semel, Wiig, & Secord, 2003). The CREVT-2 is a measure of children's receptive and expressive lexical knowledge, and the two CELF-4 subtests are indices of sentence-level receptive and expressive knowledge and abilities. Because some subtests were standardized with deviation quotients (M = 100, SD = 15) and others were standardized with scaled scores (M = 10, SD = 3), we converted the children's norm-referenced scores to a *z* score scale (M = 0, SD =1) representing the number of standard deviations from the mean on each subtest. A mean composite *z* score was then calculated for the three lowest scores.

The composite *z* score for all the children in the SLI group was −1.48 with a *SD* of 0.39 (range = −2.73 to −1.00). The overwhelming majority of the children in the SLI group (84.6%) had mixed receptive–expressive disorders as evidenced by performance at or below the criterion value of −1 *SD* on the expressive and receptive subtests. Some children (14.5%) exhibited expressive-only disorders, and a very small minority (1%) exhibited receptive-only disorders. With respect to the language domain, 74.4% of the children performed at or below the criterion value on subtests in both the lexical and sentential domains; 18.8% had difficulties on the grammatical subtests only, and 6.8% had difficulties on the lexical subtests only. Regarding nonverbal IQ, the children attained a mean score of 97.98 (SD = 13.80) on the four Leiter subtests (Roid & Miller, 1997).

TD Group

Children were defined as TD if their composite *z* score on the lowest three language measures was above −1 *SD* from the mean. The average composite *z* score for the qualifying measures for the children in the TD group was 0.08 with a *SD* of 0.60 (range = −0.96 to 1.89), which was significantly larger than that for the SLI group, $F(1, 233)$ = 556.74, $p < .0001$, $\eta^2 = .71$. Although the children in both the TD and SLI groups exhibited normal-range nonverbal IQ, the children in the TD group obtained a significantly higher score than the children in the SLI group, $F(1, 233)$ = 46.22, $p < .0001$, $\eta^2 = .17$, a finding that is consistent across the SLI literature (Gallinat & Spaulding, 2014). Johnston (1982) and Swisher, Plante, and Lowell (1994) suggested that nonverbal IQ scores may not serve as measures of general intelligence in children with language impairment. However, performance on these measures may represent the degree of multiple nonlinguistic deficits. Following the recommendations of Swisher et al., we did not use nonverbal IQ as a matching variable. Instead, it was used as a covariate in all between-groups analyses to control for any potential influence that nonlinguistic deficits might have on the children's language. Separate analyses of covariance (ANCOVAs, with nonverbal IQ as the covariate) were performed to verify that, relative to the SLI group, the TD group obtained significantly higher standard scores on each entrance language test. With respect to the language scores, results indicated that the TD group attained a significantly higher score on each measure: CREVT-2 receptive, $F(1, 233) = 61.85$, $p < .0001$, $\eta^2 = .21$; CREVT-2 expressive, $F(1, 233) = 37.31$, $p < .0001$, $\eta^2 = .14$; CELF-4 concepts and following directions, $F(1, 233) = 50.29$, $p < .0001$, $\eta^2 = .18$; and CELF-4 recalling sentences, $F(1, 233) = 63.30$, $p < .0001$, $\eta^2 = .21$. Table 1 presents summary language and IQ scores for the two groups, and Table 2 provides a breakdown of scores by age group by subject group.

To avoid selection bias and distortion of the results due to differences in participant enrollment, propensity score matching was used to create the SLI and TD groups from a larger pool of 383 children (127 SLI, 256 TD) who completed the project.[1] Using multivariate logistic regression, a propensity score was calculated for each of the 383 children in the complete participant pool on the basis of the moderating variables of age (continuous variable), gender (dichotomous variable: M or F), mother's education level (dichotomous variable: no college degree [high school, some college but no degree] vs. college degree [associate, bachelor's, master's, or doctorate]), and family income (dichotomous variable: annual income less than \$30,000 vs. annual income greater than \$30,000). Mother's education and family income were used as proxy measures of socioeconomic status (Shavers, 2007). Individual children with SLI and TD children were then matched on their propensity scores, yielding samples of 117 children in each group. Demographic data for the two groups are presented in Table 3. Subsequent nonparametric analyses revealed that the groups were not significantly different with respect to age, gender, mother's education, or family income.

To examine developmental changes in children's sentence comprehension, children in each group were divided into two age bands (Montgomery, Evans, Gillam, Sergeev, & Finney, 2016). The younger band had a mean age of 8:1 years (7:0–9:3), and the older had a mean age of 10:8 years (9:4–11:11). These age bands were motivated on the findings of Montgomery et al. (2016), who observed age-related improvement in a large sample of TD children ($N = 256$) for the same sentences used here.

To better understand the comprehension of both the children with SLI and the TD children, 40 healthy English-speaking undergraduate and graduate students (age_M = 23 years) completed the comprehension task. Each participant reported typical speech, language, and academic development and no history of psychiatric difficulties. Each participant also passed a hearing screening.

Sentence Comprehension Task

Children's comprehension of canonical (SVO, SR) and noncanonical (passive, OR) structures was assessed using our "Whatdunit" agent selection task (Montgomery et al., 2016).[2] Children were told that they would hear a man saying some funny sounding sentences about one thing doing an action on another thing. They were told that after each sentence three pictures would appear at the bottom of the computer screen and to touch the picture of the "thing that did the action" as quickly as they could, thereby indexing their identification of the agent.

Sentence Stimuli

The canonical structures included 33 SVOs (*The square had changed the bed under the very new dry key*) and 33 center-embedded SRs in which the head/first NP (NP1) functioned as the subject in the relative clause (*The watch that had hugged the truck behind the kite was bright*). The noncanonical structures included 33 verbal *be* passives (*The watch was bumped by the wheel near the very bright clock*) and 33 ORs in which NP1 functioned as object/patient in the relative clause (*The chair that the bread had splashed under the square was new*). The Appendix displays example sentences.

Each sentence was reversible and contained 12 words. Each sentence also included a prepositional phrase following the second NP in which a third noun appeared. Including a prepositional phrase permitted controlling sentence length across sentence types without altering fundamental syntactic form. For the SRs and ORs, the only action verb appeared in the embedded relative clause. The intent of this control was to circumvent potentially increasing difficulty associated with processing two action verbs, one in the embedded clause and one in the main clause. Verb tense in all of the sentences was past tense, with SVOs, SRs, and ORs being past perfect tense (Love, 2007). For the ORs, the relative pronoun *that* always appeared as its presence appears to enhance comprehension relative to when it is absent (Hakes, Evans, & Brannon, 1976). The 132 sentences were

[1]A propensity score is the conditional probability of a child being enrolled in the SLI or control (TD) group given his or her key baseline characteristics (in our case, age, gender, mother's education, family income). Due to its ability to match groups on a high-dimensional set of characteristics—that is, simultaneous matching on several categorical and continuous variables—propensity score technique has become a critical statistical method in modern clinical research (D'Agostino, 1998; Rosenbaum & Rubin, 1983).

[2]Note. From Whatdunit? Developmental changes in children's syntactically based sentence interpretation abilities and sensitivity to word order by J. Montgomery, J. Evans, R. Gillam, A. Sergeev, and M. Finney, 2016, *Applied Psycholinguistics, 37,* p. 1281. Copyright 2016 by Cambridge University Press. Reprinted with permission.

Table 1. Mean (*M*) standard scores and standard deviations (*SD*) on the norm-referenced test measures administered to the children with specific language impairment (SLI) and typically developing (TD) children.

Measure	SLI (*N* = 117)	TD (*N* = 117)	Cohen's *d*
Nonverbal IQ			
Leiter[a]			
M	98	110	–0.77
SD	13	14	
Range	76–139	76–141	
Lexical			
CREVT-2, Receptive[b]			
M	87	105	–1.22
SD	9	11	
Range	62–112	81–146	
CREVT-2, Expressive[c]			
M	81	101	–1.32
SD	10	12	
Range	54–101	69–134	
Sentential			
CELF-4, Concepts & Direct[d]			
M	6	11	–1.33
SD	3	2	
Range	1–13	6–15	
CELF-4, Recalling Sent[e]			
M	5	10	–1.51
SD	2	2	
Range	1–11	4–18	
Qualifying *z* score[f]			
M	–1.49	0.08	–3.10
SD	0.39	0.60	
Range	–2.73 to –1.00	–0.96 to 1.89	

Note. CREVT-2 = Comprehensive Expressive-Receptive Vocabulary Test–Second Edition; CELF-4 = Clinical Evaluation of Language Fundamentals–Fourth Edition.

[a]Average score on four nonverbal subtests (figure ground, form completion, sequential order, and repeated patterns) from the visualization and reasoning battery of the Leiter International Performance Scale–Revised: *M* = 100, *SD* = 15. [b]*M* = 100, *SD* = 15. [c]*M* = 100, *SD* = 15. [d]CELF-4 Concepts & Directions: *M* = 10, *SD* = 3. [e]CELF-4 Recalling Sentences: *M* = 10, *SD* = 3. [f]Average *z* score on the three lowest lexical and sentential measures.

arranged into three blocks of 44 items. One block was presented during each of the three testing sessions.

The sentences were created using a pool of 33 nouns, 22 verbs, and three prepositions. Sentences were constructed to be semantically implausible and to express highly improbable events, ensuring the children would rely just on word order cues to guide comprehension. Semantic implausibility was created in two related ways. First, noun animacy was controlled by selecting inanimate/object nouns as the agent and patient of the sentences. Second, we violated typical predicate argument structure—that is, verb selection restriction rules. Verb selection restriction rules are constraints on verbs that determine what semantically appropriate noun arguments a verb can take (Altmann & Kamide, 1999; Ferretti, McRae, & Hatherell, 2001). For example, in the sentence *The girl is eating the cookie* the verb *eating* specifies that the subject/agent noun must be animate and the object/patient noun be an edible. In the case of a verb selection violation, one or more nouns are semantically inappropriate (*The girl is riding the cookie*). At the same time, a verb selection violation gives rise to a violation of the natural affordance between the nouns in a sentence. In most sentences, an affordance between the nouns and their associated semantic roles is expressed. Affordance refers to the ways in which people interact with objects in the world with the interaction reflecting intrinsic constraints that occur between the entities (Gibson, 1979; Glenberg et al., 2009). In *The girl is eating the cookie,* the affordance between the two nouns (*girl, cookie*) is a natural one as encoded/expressed through the verb *eating*. Sentences involving natural affordance also typically express probable events. Such event or pragmatic knowledge is used by adults (Matsuki et al., 2011), TD children (Chapman & Kohn, 1978; Friedrich & Friederici, 2005; Pereyra, Klarman, Lin, & Kuhl, 2005), and children with SLI (Evans, 2002; Evans & MacWhinney, 1999) to guide sentence comprehension.

Two other findings in the psycholinguistic literature motivated our construction decisions. First, it has been shown that agent nouns and patient nouns can prime the activation of verbs (McRae, Hare, Elman, & Ferretti, 2005). Second, there is some evidence (Kamide, Altmann, & Haywood, 2003) that already activated arguments can prime the activation of other arguments appearing later in a sentence. By using inanimate nouns and violating predicate

Table 2. Mean (*M*) and standard deviations (*SD*) on the standardized tests administered to the younger and older children with specific language impairment (SLI) and typically developing (TD) children.

	Younger groups			Older groups		
Measures	SLI	TD	Cohen's *d*	SLI	TD	Cohen's *d*
Nonverbal IQ						
Leiter[a]						
M	99	110	−0.76	96	110	−0.99
SD	14	14		13	13	
Range	76–139	76–141		76–127	76–133	
Lexical						
CREVT-2, Receptive[b]						
M	86	104	−1.83	88	106	−1.73
SD	7	12		9	11	
Range	76–109	81–146		62–112	84–124	
CREVT-2, Expressive[c]						
M	81	102	−1.86	80	100	−1.74
SD	9	13		11	11	
Range	60–101	69–134		54–99	72–126	
Sentential						
CELF-4, Concepts & Direct[d]						
M	6	11	−1.85	5.4	11	−2.14
SD	2	2		3.2	2	
Range	1–13	6–15		1–13	6–14	
CELF-4, Recalling Sent[e]						
M	5	11	−2.41	5.7	10	−2.31
SD	2	2		2.0	2	
Range	1–11	4–18		1–11	6–17	
Qualifying *z* score[f]						
M	−1.5	0.1	−3.92	−1.5	0.1	−3.13
SD	0.4	0.4		0.43	0.6	
Range	−2.29 to −1.0	−0.95 to 1.89		−2.73 to −1.0	−0.95 to 1.09	

Note. CREVT-2 = Comprehensive Expressive-Receptive Vocabulary Test–Second Edition; CELF-4 = Clinical Evaluation of Language Fundamentals–Fourth Edition.

[a]Average score on four nonverbal subtests (figure ground, form completion, sequential order, and repeated patterns) from the visualization and reasoning battery of the Leiter International Performance Scale–Revised: *M* = 100, *SD* = 15. [b]*M* = 100, *SD* = 15. [c]*M* = 100, *SD* = 15. [d]CELF-4 Concepts & Directions: *M* = 10, *SD* = 3. [e]CELF-4 Recalling Sentences: *M* = 10, *SD* = 3. [f]Average *z* score on the three lowest lexical and sentential measures.

argument expectancies, we greatly reduced possible priming effects in the sentences. Together, such manipulations rendered all of the sentences implausible, expressing highly improbable events, thereby severely limiting the availability of semantic–pragmatic cues to help the children determine which noun functioned as agent and which as patient.

To minimize the influence that lexical knowledge may play in children's comprehension (Borovsky et al., 2012), we carefully controlled the properties of the nouns, verbs, and prepositions appearing in the sentences. This control was essential given the documented lexical deficits of children with SLI (e.g., Mainela-Arnold, Evans, & Coady, 2008, 2010; McGregor, Newman, Reilly, & Capone, 2002). All of the main lexical items had spoken word frequency ratings of 6:0 years or less (Moe, Hopkins, & Rush, 1982) and age of acquisition ratings of 3.6 years or younger (Cortese & Khanna, 2008). The nouns also had high imageability (> 500), concreteness (> 500), and familiarity (> 500) ratings (Coltheart, 1981; Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012; Vitevitch & Luce, 2004). A final confirmation that the nouns in the sentences were well known by the children comes from the fact that each subject group performed at ceiling (SLI = 98% correct, TD = 99% correct) on a rapid automatic naming task comprising the 33 sentence nouns. Images of the nouns were presented in random order on the computer, and the children were instructed to name each picture as quickly as they could. This task was administered as part of the larger ongoing project.

Each of the nouns appeared with equal frequency as NP1, NP2, and NP3 across the sentences, and the verbs and prepositions occurred with equal frequency. The images corresponding to the nouns were color drawings of simple objects (e.g., bed, coat, spoon) standardized for name and image agreement, familiarity, and visual complexity (Rossion & Pourtois, 2004). Last, the sentences were recorded at a normal speaking rate (approximately 4.4 syllables/s; Ellis Weismer & Hesketh, 1993) and with normal prosodic variation by an adult male speaker of Midwestern American English. All audio files were digitized (44 kHz), low-pass filtered (20 kHz), and normalized for intensity.

The "Whatdunit" task has strong validity and reliability (Montgomery et al., 2016). On the basis of a large

Table 3. Participant demographics for the children with specific language impairment (SLI) and typically developing (TD) children.

Demographic	SLI (*N* = 117)	TD (*N* = 117)
Age_M (years:months)	9:5	9:5
Gender, %		
Male	57	63
Female	43	36
Race and ethnicity, %		
White (not Hispanic)	61	72
African American	10	0
Hispanic	12	12
Asian	4	4
American Indian, Native Hawaiian	3	3
More than one race	10	9
Mother's education, %		
No response	1	1
High school degree	20	16
Some college	30	27
Associate degree	17	11
Bachelor's degree	24	23
Graduate degree	6	20
Family income, %		
\$0–\$25,000	42	32
\$26,000–\$50,000	21	22
\$51,000–\$75,000	16	15
> \$75,000	21	31

sample (*N* = 256) of 7- to 11-year-old TD children, internal construct validity of the canonical structures was .84 and .89 for the noncanonical structures. The correlations between the individual canonical and noncanonical items ranged between .31 and .35. Such results support the construct of canonical/noncanonical differences. Concurrent validity was likewise very good as evidenced by strong correlations between overall score on the "Whatdunit" task and lexical measures (CREVT-2, .62), sentence-level measures (CELF-4 subtests, .60), and overall language (Test of Narrative Language Narrative Language Ability Index, .64). Internal consistency for each sentence type was very strong (SVO = .88, SR = .86, passive = .95, OR = .94) and for the task as a whole (.97).

Procedures

The children were seen individually in a quiet laboratory over three testing sessions, each lasting about 2 hr, including rest breaks. To record the accuracy of the children's responses as well as ensure a fixed random order of presentation of the items and position of the correct answer, delivery of the stimuli was controlled using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) running on a laboratory laptop connected to a 17-in. Elo Touch Screen monitor (Elo Touch Solutions, model 1715L).

Children sat at a table in front of a touch screen. Children placed their arm in a comfortable position on the table so that the fingers of their dominant hand rested on a red dot located in the center of the bottom edge of the monitor, just below the touch screen. Children were instructed to leave their fingers on the dot until they were ready to touch the screen. Stimuli were presented binaurally under noise-canceling headphones at a comfortable listening level determined by the child. Before receiving the experimental items, all participants demonstrated that they understood the task during training on sentences with and without event probability cues.

Data Analysis

A three-way mixed ANCOVA was conducted with subject group (SLI, TD) as the between-subjects variable and age group (younger, older) and sentence type (SVO, SR, passive, OR) as within-subject variables. Because the groups significantly differed in nonverbal IQ, the children's total raw score on the four Leiter measures was used as the covariate in all between-groups analyses. The percentage accuracy scores were not normally distributed across groups and sentence types, thus they were converted to *z* scores, which yielded distributions more closely approximating normality (skewness < ± 2.1; West, Finch, & Curran, 1996). The converted scores were used in the parametric analyses. Pairwise comparisons were conducted to examine the nature of any significant interactions with a least significant difference α level of .05. Percentage scores are reported in the text and Table 4 for ease of interpretation.

Table 4. Mean comprehension accuracy (percentage correct) by subject group (specific language impairment [SLI], typically developing [TD]), age group, and sentence type (subject–verb–object [SVO], subject relative [SR], passive, object relative [OR]).

	Sentence Type			
Group	SVO	SR	Passive	OR
SLI				
Younger (*n* = 53)				
M	55	56	29	30
SD	25	21	20	13
Range	6–94	15–94	0–79	9–69
Older (*n* = 64)				
M	72	69	33	27
SD	21	22	23	17
Range	16–100	21–100	0–97	0–73
Grand (*N* = 117)				
M	65	63	30	28
SD	24	22	22	15
Range	6–100	15–100	0–97	0–73
TD				
Younger (*n* = 61)				
M	77	74	49	39
SD	20	20	29	26
Range	6–100	27–100	0–100	0–100
Older (*n* = 56)				
M	87	82	62	57
SD	13	14	31	30
Range	52–100	42–100	0–100	3–97
Grand (*N* =117)				
M	83	78	55	48
SD	18	18	30	29
Range	6–100	27–100	0–100	0–100
Adults (*N* = 40)				
M	98	97	88	86
SD	5	6	27	28
Range	82–100	82–100	0–100	0–100

Results

Sentence Comprehension Accuracy

Results of the omnibus ANCOVA revealed a significant subject group × age group × sentence type interaction, $F(1, 229) = 7.23$, $p < .0001$, $\eta^2 = .03$. To derive maximally interpretable results, the interaction was broken down into two separate subject group × age group × sentence type ANCOVAs: one focusing on canonical sentences (SVO, SR) and the other on noncanonical sentences (passive, OR).

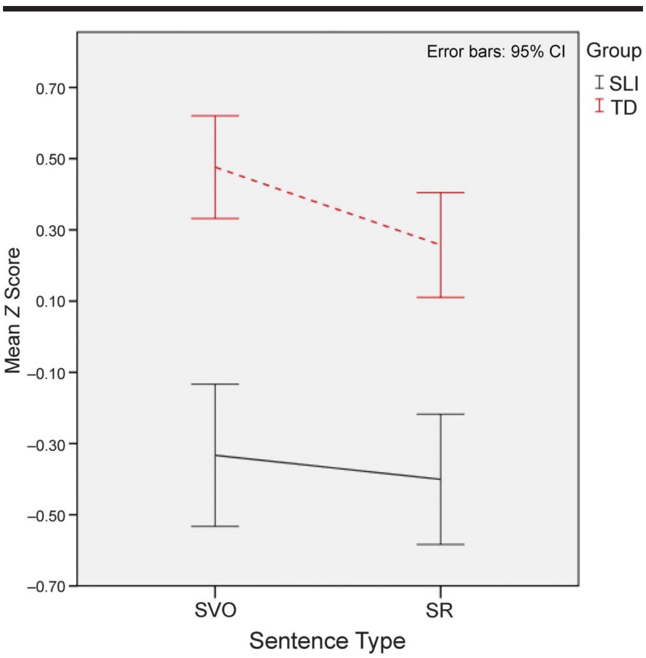
Canonical Structures

Unlike the omnibus test, the subsequent three-way subject group × age group × canonical sentence type interaction was not significant. However, the subject group × sentence type interaction was significant, $F(1, 229) = 4.05$, $p = .05$, $\eta^2 = .02$, as was the age group × sentence type interaction, $F(1, 229) = 4.64$, $p = .03$, $\eta^2 = .02$.

Subject group × sentence type interaction. Between subject group pairwise comparisons revealed significantly poorer comprehension in the SLI group than the TD group on both the SVOs (TD = 83%, SLI = 65%, $p < .05$) and SRs (TD = 78%, SLI = 63%, $p < .05$). A within-SLI group comparison indicated no significant difference ($p > .05$) between the SVOs (65%) and SRs (63%). By contrast, a within-TD group comparison showed the TD children had significantly better comprehension ($p < .05$) of SVOs (83%) than SRs (78%). This two-way interaction appears in Figure 1.

Age group × sentence type interaction. Between age group pairwise comparisons revealed that the older children outperformed the younger children on both the SVOs (older = 80%, younger = 67%, $p < .05$) and SRs (older = 75%, younger = 66%, $p < .05$). Within age group comparisons indicated that younger children were comparable in comprehending SVOs (67%) and SRs (66%; $p > .05$) whereas older children were significantly better at comprehending SVOs (80%) than SRs (75%; $p < .05$). The two-way interaction appears in Figure 2.

Figure 1. Comprehension of subject–verb–object (SVO) and subject relative (SR) sentences by subject group (specific language impairment [SLI], typically developing [TD]). CI = confidence interval.



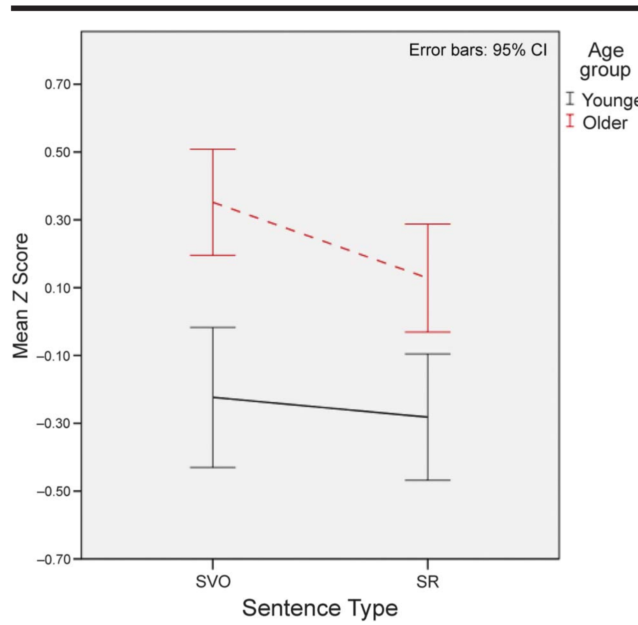
Noncanonical Structures

The subject group × age group × noncanonical sentence type ANCOVA yielded a significant three-way interaction, $F(1, 229) = 9.77$, $p = .002$, $\eta^2 = .04$. Pairwise comparisons were conducted to examine the nature of the interaction, first on the passive sentences followed by the OR sentences.

Passive sentences. Between subject group and age group pairwise comparisons revealed that younger children with SLI performed significantly worse than younger TD children (SLI = 29%, TD = 49%, $p < .05$). Older children with SLI also performed significantly worse than older TD children (SLI = 33%, TD = 62%, $p < .05$). Within the SLI group, older and younger children were not significantly different (older SLI = 33%, younger SLI = 29%, $p > .05$). However, among the TD children, older children performed significantly better than younger children (older TD = 62%, younger TD = 49%, $p < .05$).

OR sentences. Results of the between subject group and age group analysis revealed no significant difference ($p > .05$) between the younger children with SLI (30%) and younger TD children (39%). But among older children, TD children (57%) significantly outperformed ($p < .05$) the

Figure 2. Comprehension of subject–verb–object (SVO) and subject relative (SR) sentences by age group. CI = confidence interval.



children with SLI (27%). Within the SLI group, older and younger children did not differ in comprehension (older = 27%, younger = 30%, $p > .05$). However, among the TD children, older children performed significantly better than younger children (older = 57%, younger = 40%, $p < .05$). Figures 3 and 4 illustrate the subject group × age group interaction for passives and ORs, respectively. Table 4 presents summary statistics of the children's comprehension accuracy.

Comparison of Older SLI and Younger TD Groups

In keeping with Dick et al. (2004), we compared the older SLI group with the younger TD group on each sentence type. The differences between the older SLI group and the younger TD group were not significant for the two canonical structures: SVO (older SLI = 72%, younger TD = 77%), $F(1, 124) = 2.83$, $p = .09$, $\eta^2 = .02$, and SR (older SLI = 69%, younger TD = 74%), $F(1, 124) = 2.70$, $p = .10$, $\eta^2 = .02$. For the noncanonical sentences, the older SLI group performed significantly worse than the younger TD group on both passives (older SLI = 33%, younger TD = 49%), $F(1, 124) = 13.71$, $p < .0001$, $\eta^2 = .10$, and ORs (older SLI = 27%, younger TD = 39%), $F(1, 124) = 11.34$, $p < .001$, $\eta^2 = .09$.

Adult Performance

Recall that a group of adults also completed the task to provide an adult reference point against which to compare the children's performance. The adults performed at ceiling on the SVOs (98%) and SRs (97%). They also yielded very high levels of accuracy on the passives (88%) and ORs (86%).

Figure 3. Comprehension of passive sentences by subject group (specific language impairment [SLI], typically developing [TD]) and age group. CI = confidence interval.

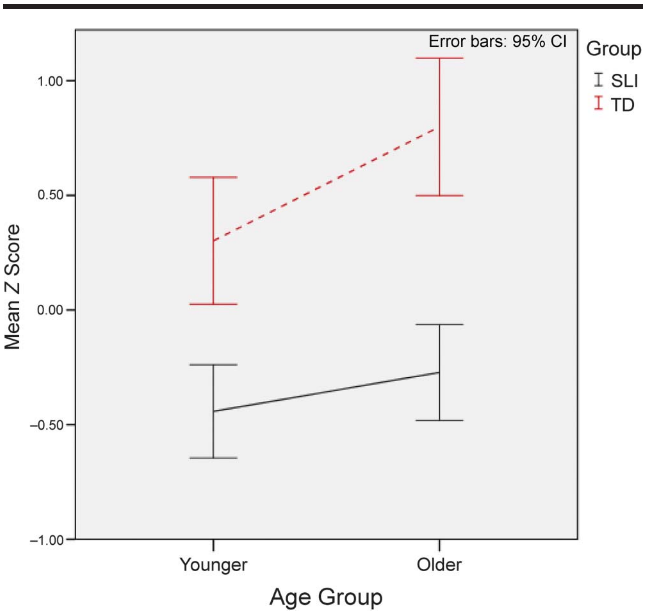
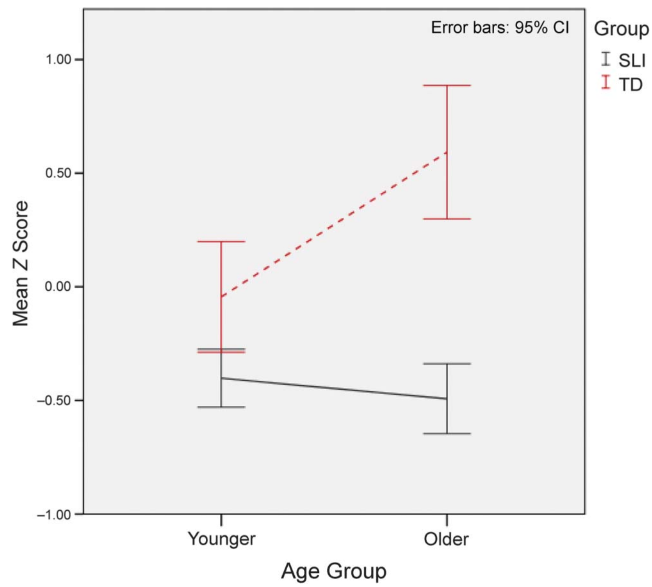


Figure 4. Comprehension of object relative sentences by subject group (specific language impairment [SLI], typically developing [TD]) and age group. CI = confidence interval.



Error Patterns

Children's error patterns were also examined. Two types of errors were possible. Object noun (object) errors reflected children's improper selection of the object noun as the agent. Prepositional phrase noun (PPN) errors reflected an improper selection of the noun in the prepositional phrase near the end of the sentence as the agent. A subject group (SLI, TD) × error type (object, PPN) ANCOVA (using nonverbal IQ as the covariate) was conducted on each of the sentence types separately (SVO, SR, passive, OR). The analyses were performed on percentage error scores as they were normally distributed (skewness < ± 2.1; West et al., 1996).

For the SVOs, the error type main effect, $F(1, 231) = 8.70$, $p = .004$, $\eta^2 = .04$, and the subject group main effect, $F(1, 231) = 16.17$, $p < .0001$, $\eta^2 = 07$, were both significant. Children made significantly more object errors than PPN errors, and the TD group performed significantly better than the SLI group. For the SRs, only a significant subject group main effect emerged, $F(1, 231) = 10.57$, $p = .0001$, $\eta^2 = .04$, which favored the TD group. For the passives, the error type effect was significant, $F(1, 231) = 6.87$, $p = .009$, $\eta^2 = .03$, with children making significantly more object errors than PPN errors. The subject group effect was also significant, $F(1, 231) = 23.31$, $p < .0001$, $\eta^2 = .09$, favoring the TD group. In a similar manner, for the ORs, there were significant main effects for error type, $F(1, 231) = 7.17$, $p = .008$, $\eta^2 = .03$, with children making significantly more object errors than PPN errors, and for subject group, $F(1, 231) = 19.55$, $p < .0001$, $\eta^2 = .08$, with TD children making fewer errors than children with SLI. Last, the adults also produced more object errors than PPN errors in both passives (10% vs. 2%) and ORs (11% vs. 3%). Table 5 displays summary error data for the SLI and TD groups.

Discussion

Comprehension Accuracy

The children with SLI evidenced significantly poorer comprehension of all sentence types than their TD peers, supporting our first prediction. On the SVOs, the SLI group achieved just 65% accuracy whereas the TD group performed with 83% accuracy. On the SRs, the SLI group performed with 63% accuracy while the TD group performed with 78% accuracy. These findings indicate that the children with SLI had reduced sensitivity and ability to use canonical word order compared with same-age peers. However, despite the lower scores of the children with SLI, it is important to point out that they performed well above chance (chance = 33%), indicating that they had some sensitivity and ability to use canonical word order to guide comprehension.

Both the SLI and TD groups performed substantially more poorly on the passives and ORs relative to the canonical structures with the SLI group performing significantly worse than the TD group. Although the groups yielded roughly comparable decrements (SLI 35% less accuracy, TD group 30% less accuracy), the groups' absolute scores were vastly different. It was striking that the SLI group performed near or just below chance on both the passives (30%) and ORs (28%) while the TD group performed well above chance (passive = 55%, OR = 48%). These findings indicate that children with SLI have substantially reduced sensitivity and ability to use noncanonical word order to guide comprehension. Additional evidence of this claim comes from the comparison between the older children with SLI (9:4–11:11) and younger TD children (7:0–9:3). Even compared with younger TD children, older children with SLI show significantly poorer comprehension of passives and ORs, findings that are consistent with Dick et al. (2004). The converse is that the older SLI and younger TD children showed comparable comprehension of SVOs and SRs—also consistent with Dick et al. (2004).

One might argue that the children with SLI sacrificed accuracy for speed given that we instructed all of the children to be as quick as possible in selecting the agent of the sentence. A speed–accuracy trade-off is unlikely, however. Both groups' yielded comparable speed of comprehension ($p > .05$) across each sentence type with the exception of the passives; the TD group was faster.

We refer to the performance of the adults who completed the same task to further contextualize our findings. Recall that the adults' performance on the SVOs and SRs was at ceiling (98%, 97%) and on the passives and ORs, their performance was very good (88%, 86%, respectively). Such strong performance by the adults is expected given that English is a word order language, and adults know to attend to word order cues over animacy cues (MacWhinney, Bates, & Kliegl, 1984). Such results show, first, that even school-age TD children have yet to acquire adult-like sensitivity to and use of noncanonical or canonical word order cues. Second, the results further accentuate the severe limitations that children with SLI have attending to and using word order cues, especially noncanonical cues. Third, the developmental trajectory in learning and using both canonical and noncanonical word orders by children with SLI is extremely protracted.

The special difficulties posed by noncanonical structures for children with SLI is consistent with the syntax-specific deficit hypothesis of van der Lely and colleagues (Marinis & van der Lely, 2007; Marshall et al., 2007). However, it is important to point out that this study was not designed to test the theoretical merits of the hypothesis. All that can be said is that the children predictably had substantial difficulties with noncanonical structures. The poor comprehension of the canonical structures by the children, however, is not in keeping with the hypothesis.

Table 5. Mean percentage comprehension errors (object error, prepositional phrase noun [PPN] error) by subject group (specific language impairment [SLI], typically developing [TD]) and sentence type (subject–verb–object [SVO], subject relative [SR], passive, object relative [OR]).

	Sentence Type				
Group	SVO	SR	Passive	OR	Grand
SLI					
Object error					
M	16	19	50	52	34
SD	9	10	25	22	17
Range	0–36	0–50	0–100	12–100	0–100
PPN error					
M	19	18	19	19	19
SD	21	16	21	16	18
Range	0–88	0–73	0–85	0–58	0–88
TD					
Object error					
M	10	12	38	43	26
SD	9	10	29	27	19
Range	0–36	0–40	0–100	0–97	0–100
PPN error					
M	6	10	7	10	8
SD	11	10	11	11	11
Range	0–82	0–40	0–79	0–67	0–82

Developmental Changes in Sentence Comprehension and Sensitivity to Word Order

The children with SLI and TD peers revealed different developmental trajectories in their use of word order cues. Both groups exhibited age-related improvement in comprehension of SVOs and SRs, supporting our second prediction. These findings indicate that, despite their overall poorer performance, children with SLI, like their same-age peers, continue to develop sensitivity to and facility for using high-frequency canonical word orders (Dick et al., 2004).

The converse is that the picture was very different for the noncanonical structures. As predicted, the TD children revealed developmental improvement in the comprehension of passives and ORs (Dick et al., 2004). However, the children with SLI exhibited no age-related improvement for either passives or ORs, contrary to our prediction and the findings of Dick et al. (2004). In fact, the absolute scores of our older and younger children with SLI were essentially identical. It is interesting to note that, regarding the ORs, the younger children with SLI and younger TD children yielded comparable comprehension. These findings were unexpected and suggest that OR word order is difficult even for TD children at a young age. But the TD children exhibited developmental growth, and by 9:4–11:11 were significantly outperforming their peers with SLI. The older children with SLI, it is striking to note, continued to perform like their younger SLI counterparts, indicating that through age 11 years these children show remarkable stability (i.e., lack of growth) in their appreciation and use of OR word order. The absence of an age effect for passives was somewhat surprising, though. Passives (*The train was watched by the bed behind the very cold cake*), unlike ORs (*The train that the bed had watched near the boot was dry*), contain (morpho)syntactic cues, such as the past participle *-ed* coupled with the auxiliary verb *was* (*was watched*) and an adjunct prepositional *by* phrase (*by the bed*). That the children with SLI revealed no developmental effect implies a significant limitation in learning that such cues represent markers of a passive. One might argue, however, that difficulty processing the brief-duration inflection *-ed* (e.g., Leonard, 2014; Montgomery & Leonard, 2006), which also appeared in all of the sentences, may have led to the children's poor comprehension. Even if this were the case, the presence of the *by* phrase lent the children a salient syntactic cue. Overall, then, we would argue that the children with SLI showed developmental stagnation in the learning of both noncanonical word order structures.

Why Sentence Comprehension Is So Difficult in the Absence of Semantic Cues

Understanding a sentence involves building syntactic structure and assigning semantic meaning to structure. Findings from the adult literature make it clear that sentence comprehension is driven by the predictive relationship between verbs and their syntactic arguments as well as by pragmatic knowledge (Altmann & Kamide, 1999; Kuperberg, 2007; Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Paczynski & Kuperberg, 2012). Emerging evidence in the developmental literature shows that young children also use all available cues to engage in predictive processing during comprehension (Borovsky et al., 2012). In the present study, all of the sentences, by design, were absent semantic–pragmatic cues as all of the nouns were inanimate and the availability of usual predicate argument expectancies and potential verb/noun priming effects were controlled. The absence of these cues obviously made comprehension more challenging for the children and disproportionately so for those with SLI.

A recent and influential framework of sentence comprehension from the adult literature supplies additional evidence of the importance of linguistic cues and, more critical, distinctive cues, supporting comprehension. The framework centers on the role of item retrieval—to be specific, similarity-based retrieval interference—as a critical determinant of complex sentence comprehension (Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; Lewis, Vasishth, & Van Dyke, 2006; Van Dyke, 2007; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006, 2011). Take the OR sentence *The train that the bed had watched near the boot was dry*; the comprehender must momentarily store both NP1 (*the train*) and NP2 (*the bed*) until processing the embedded VP/syntactic gap (*had watched*). At this point, the comprehender must selectively reactivate NP1 over NP2 to establish a filler–gap dependency (…*the bed had watched the train*…) and assign the role of agent to NP2 and patient to NP1, enabling him or her to understand who did what to whom. According to the framework, if the cues available at retrieval are not sufficiently discriminating to reactivate the target (NP1) over other competing semantically and/or syntactically similar items in memory (NP2 and any other NPs), then retrieval interference arises. Under such circumstances, an inappropriate NP may be retrieved, and comprehension is compromised. It is important to note, though, that when distinctive cues are available, retrieval interference is reduced or even eliminated (Van Dyke & McElree, 2011). On this view, we would argue that the children with SLI were much more vulnerable than the TD children to similarity-based retrieval interference.

This framework also would appear to hold for the children's poor comprehension of the passives. In passives (*The train was watched by the bed behind the very cold cake*), the verb (*watched*) cues the reactivation of NP1. It is important to note, though, that there is evidence in the adult literature showing that NP1 is not immediately reactivated after the verb but up to 1 s after the verb (Osterhout & Swinney, 1993), which, in our case, would likely be sometime after NP2 (*The train was watched by the bed* [NP1 reactivation] *behind the very cold cake*). Assuming this is the case, the children would need to hold both NP1 and NP2 in memory until NP1 needs to be reactivated, which would lead to competition between the NPs as to which should be selectively reactivated. This possibility is bolstered on two grounds. First, our own adult participants performed markedly lower on the passives (and ORs) than on the SVOs and SRs. Second, even adults are less accurate in assigning proper semantic roles in passive sentences when both NP1 and NP2 are animate than when there is an animacy mismatch between NP1 and NP2—that is, NP1 is inanimate and NP2 is animate (Ferreira, 2003). In a similar manner, adults show slower processing of passive sentences when the predicate argument structure is less predictable than when it is more predictable (Ambridge, Bidgood, Pine, Rowland, & Freudenthal, 2015). Given the agent and patient NPs in both noncanonical sentence types were inanimate and not usual partners of the verbs, it is no surprise that the children experienced such hardship making proper role assignments, assuming NP1 reactivation occurred.

Although we cannot make the case that the poor SVO and SR comprehension of the children with SLI was related to NP competition, it certainly is the case that the children found it difficult to make correct semantic role assignments in the absence of usual predicate argument cues. Results from a recent eye-tracking study with adolescents with SLI appear to be relevant to this interpretation (Borovsky, Burn, Elman, & Evans, 2013). Children listened to simple SVOs (*The dog hides the X*) while looking at four noun pictures, one that was the target noun argument (*bone*) of the verb, one that was a related noun argument to the agent noun (*cat*), one that was a related noun argument to the verb (*treasure*), and one that was unrelated (*ship*). Although the SLI and TD groups did not differ in fixating on the target nouns, the SLI group exhibited increased fixations to the verb-related noun images. Results were interpreted to suggest that the SLI group had lexical integration deficits. However, the results may also imply that the SLI group was not as facile as same-age peers to build an "expected" predicate argument structure (e.g., Andreu, Sanz-Torrent, & Guardia-Olmos, 2012; Pijnacker et al., 2017; Thordardottir & Ellis Weismer, 2002). Even within the adult literature, it has been shown that SVOs that violate predicate argument expectations are more poorly comprehended than SVOs expressing usual expectations (Saffran, Schwartz, & Linebarger, 1998).

Comprehension Error Patterns

Children's error patterns were examined to determine if they made more object errors than PPN errors (i.e., selection of noun appearing in prepositional phrase near the end of the sentence) for miscomprehended sentences. We anticipated that both groups would make more object errors than PPN errors in each sentence type because object errors are linguistically based. Our expectation was confirmed with one exception. For the SRs, the children showed no difference in error type. Overall, these findings suggest that the children's errors were linguistically motivated. They selected the noun from within either the main clause or the embedded clause as the agent of the sentence, not the noun appearing in a prepositional phrase.

Clinical Implications

Two immediate clinical implications arise from the present study to enhance SLI learning of noncanonical word order sentences. First, clinicians should consider starting complex sentence comprehension training with model sentences that are heavily, semantically, pragmatically loaded, marking agency with an animate noun and patient with an inanimate noun. Enriching sentences with strong animacy–inanimacy cues could provide the children with salient predicate argument expectancies, providing them a potentially reliable semantic–pragmatic bootstrap to word order learning. Second, model sentences should incorporate familiar lexical items to help keep the focus of learning on word order.

Conclusions

This study revealed three major findings. First, school-age children with SLI are significantly poorer than their TD peers at comprehending canonical and noncanonical sentences that require syntactic processing and use of only word order cues. Second, like TD peers, children with SLI demonstrate developmental improvement in the comprehension of canonical sentences but lag behind their TD peers in the acquisition of canonical word order forms. The converse is that their comprehension of noncanonical word order structures shows no improvement with age, indicating developmental stagnation and a severe limitation in their ability to glean and acquire such word order patterns relative to same-age peers. Third, similar to their TD peers, when children with SLI miscomprehend a sentence, their errors are linguistically driven.

Future studies could begin to examine in a stepwise fashion, similar to the approach used in the adult literature, the influence of semantic and pragmatic cues on the sentence comprehension of children with SLI. Such efforts could systematically introduce different cues and cue combinations with the intent of determining if there is a point at which children with SLI approach the performance of same-age peers. Also, future studies could investigate various memory-related abilities to determine their influence on the comprehension of these children. And last, future studies could investigate how children with SLI approach the fundamentally different activities of offline and online sentence comprehension, again with an eye toward understanding the linguistic, nonlinguistic, and memory-related influences on comprehension. In the aggregate, results from such efforts could shed critically important and new insights into the sentence comprehension abilities of these children, and these could inform our assessment and intervention practices.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC010883, awarded to James W. Montgomery, Ronald B. Gillam, and Julia L. Evans. We express our gratitude to all the children and their parents who participated in this study. We also thank Beula Magimairaj, Naveen Nagaraj, Misha Finney, Yazmin Ahmad Rusli, Jenny Boyden, Andrea Fung, Kelly Rogers, Llely Duarte, Katie Squires, and Allison Hancock for their invaluable assistance during various phases of this study.

References

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73,* 247–264.

Ambridge, B., Bidgood, A., Pine, J. M., Rowland, C. F., & Freudenthal, D. (2015). Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies. *Cognitive Science, 40,* 1435–1459.

American National Standards Institute. (1997). *Specifications of audiometers* (ANSI/ANS 8.3-1997, R2003). New York, NY: Author.

Andreu, L., Sanz-Torrent, M., & Guardia-Olmos, J. (2012). Auditory word recognition of nouns and verbs in children with specific language impairment (SLI). *Journal of Communication Disorders, 45,* 20–34.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). Hillsdale, NJ: Erlbaum.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 3–73). Cambridge, United Kingdom: Cambridge University Press.

Bedore, L., Peña, E., Summers, C., Boerger, K., Resendiz, M., Greene, K., Bohman, T., & Gillam, R. (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism: Language and Cognition, 15,* 619–629.

Borovsky, A., Burns, E., Elman, J., & Evans, J. (2013). Lexical activation during sentence comprehension in adolescents with history of specific language impairment. *Journal of Communication Disorders, 46,* 413–427.

Borovsky, A., Elman, J., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112,* 417–436.

Chapman, R., & Kohn, L. (1978). Comprehension strategies in two- and three-year olds: Animate agents or probable events? *Journal of Speech and Hearing Research, 21,* 746–761.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 33*(A), 497–505.

Cortese, M., & Khanna, M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods, 40,* 791–794.

D'Agostino, R. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine, 17,* 2265–2281.

Dick, F., Wulfeck, B., Krupa-Kwiatkowski, M., & Bates, L. (2004). The development of complex sentence interpretation in typically developing children compared with children with specific language impairment or early unilateral focal lesions. *Developmental Science, 7,* 360–377.

Ellis Weismer, S., & Hesketh, L. (1993). The influence of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech and Hearing Research, 36,* 1013–1025.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14,* 179–211.

Evans, J. (2002). Variability in comprehension strategy use in children with specific language impairments: A dynamical systems account. *International Journal of Language & Communication Disorders, 37,* 95–116.

Evans, J., & MacWhinney, B. (1999). Sentence processing strategies in children with expressive and expressive-receptive specific language impairments. *International Journal of Language & Communication Disorders, 34,* 117–134.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47,* 164–203.

Ferretti, R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44,* 516–547.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language, 31,* 661–681.

Friedmann, N., & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain and Language, 101,* 50–63.

Friedrich, M., & Friederici, A. (2005). Semantic sentence processing reflected in the event-related potentials of one- and two-year old children. *NeuroReport, 16,* 1801–1804.

Gallinat, E., & Spaulding, T. (2014). Difference in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 57,* 1363–1382.

Gibson, J. (1979). *The ecological approach to visual perception*. New York, NY: Houghton Mifflin.

Glenberg, A., Becker, R., Klotzer, S., Kolanko, L., Muller, S., & Rinck, M. (2009). Episodic affordances contribute to language comprehension. *Language and Cognition, 1,* 113–135.

Gordon, P., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1411–1423.

Gordon, P., Hendrick, R., & Levine, W. (2002). Memory load interference in syntactic processing. *Psychological Science, 13,* 425–430.

Haegeman, L. (1994). *Introduction to government and binding theory* (3rd ed.). Oxford, United Kingdom: Blackwell.

Hakes, D., Evans, J., & Brannon, L. (1976). Understanding sentences with relative clauses. *Memory & Cognition, 4,* 283–290.

Hestvik, A., Schwartz, R., & Tornyova, L. (2010). Relative clause gap-filling in children with specific language impairment. *Journal of Psycholinguistic Research, 39,* 443–456.

Johnston, J. R. (1982). Interpreting the Leiter IQ: Performance profile of young normal and language-disordered children. *Journal of Speech and Hearing Research, 25,* 291–296.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49,* 133–159.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146,* 23–49.

Kuperberg, G., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes, 21,* 489–530.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, Instruments, & Computers, 44,* 978–990.

Leonard, L. (2014). *Children with specific language impairment* (2nd ed.). Cambridge, MA: MIT Press.

Leonard, L., Deevy, P., Fey, M., & Bredin-Oja, S. (2013). Sentence comprehension in specific language impairment: A task designed to distinguish between cognitive capacity and syntactic complexity. *Journal of Speech, Language, and Hearing Research, 56,* 577–589.

Lewis, R., Vasishth, S., & Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10,* 447–454.

Love, T. (2007). The processing of noncanonically ordered constituents in long distance dependencies by pre-school children: A real-time investigation. *Journal of Psycholinguistic Research, 36,* 191–207.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior, 23,* 127–150.

Mainela-Arnold, E., Evans, J., & Coady, J. (2008). Lexical representations in children with SLI: Evidence from a frequency-manipulated gating task. *Journal of Speech, Language, and Hearing Research, 51,* 381–393.
Mainela-Arnold, E., Evans, J., & Coady, J. (2010). Explaining lexical-semantic deficits in specific language impairment: The role of phonological similarity, phonological working memory, and lexical competition. *Journal of Speech, Language, and Hearing Research, 53,* 1742–1756.
Marinis, T., & van der Lely, H. (2007). On-line processing of wh-questions in children with G-SLI and typically developing children. *International Journal of Language & Communication Disorders, 42,* 557–582.
Marshall, C., Marinis, T., & van der Lely, H. (2007). Passive verb morphology: The effect of phonotactics on passive comprehension in typically developing and grammatical-SLI children. *Lingua, 117,* 302–320.
Marslen-Wilson, W., & Tyler, L. (1980). The temporal structure of spoken language understanding. *Cognition, 25,* 71–102.
Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29–63.
Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 576–585.
Matsuki, K., Hare, M., Scheepers, C., Chow, T., Elman, J., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 913–934.
McGregor, K., Newman, R., Reilly, R., & Capone, N. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45,* 998–1014.
McRae, K., Hare, M., Elman, J., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition, 33,* 1174–1184.
Moe, A., Hopkins, C., & Rush, R. (1982). *The vocabulary of first-grade children*. Springfield, IL: Thomas.
Montgomery, J., & Evans, J. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52,* 269–288.
Montgomery, J., Evans, J., & Gillam, R. (2009). Relation of auditory attention and complex sentence comprehension in children with specific language impairment: A preliminary study. *Applied Psycholinguistics, 30,* 123–151.
Montgomery, J., Evans, J., Gillam, R., Sergeev, A., & Finney, M. (2016). Whatdunit? Developmental changes in children's syntactically based sentence interpretation abilities and sensitivity to word order. *Applied Psycholinguistics, 37,* 1281–1309.
Montgomery, J., & Leonard, L. (2006). Effects of acoustic manipulation on the real-time inflectional processing of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 49,* 1238–1256.
Newcomer, P. L., & Hammill, D. D. (2008). *Test of Language Development–Primary: Fourth Edition.* Austin, TX: Pro-Ed.
Osterhout, L., & Swinney, D. (1993). On the temporal course of gap-filling during comprehension of verbal passives. *Journal of Psycholinguistic Research, 22,* 273–286.
Paczynski, M., & Kuperberg, G. R. (2012). Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language, 67,* 426–448.
Pereyra, J., Klarman, L., Lin, L., & Kuhl, P. (2005). Sentence processing in 30-month-old children: An event-related potential study. *NeuroReport, 16,* 645–648.
Pijnacker, J., Davids, N., van Weerdenburg, M., Verhoeven, L., Knoors, H., & van Alphen, P. (2017). Semantic processing of sentence in preschoolers with specific language impairment: Evidence from the N400 effect. *Journal of Speech, Language, and Hearing Research, 60,* 627–639.
Robertson, E., & Joanisse, M. (2010). Spoken sentence comprehension in children with dyslexia and language impairment: The roles of syntax and working memory. *Applied Psycholinguistics, 31,* 141–165.
Roid, G., & Miller, L. (1997). *Leiter International Performance Scale–Revised*. Wood Dale, IL: Stoelting.
Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70,* 41–55.
Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception, 33,* 217–236.
Saffran, E., Schwartz, M., & Linebarger, M. (1998). Semantic influences on thematic role assignment: Evidence from normal and aphasics. *Brain and Language, 62,* 255–297.
Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
Scott, C. (2009). A case for the sentence in reading comprehension. *Language, Speech, and Hearing Services in Schools, 40,* 184–191.
Semel, E., Wiig, E., & Secord, W. (2003). *Clinical Evaluation of Language Fundamentals–Fourth Edition*. San Antonio, TX: The Psychological Corporation.
Shavers, V. (2007). Measurement of socioeconomic status in health disparities research. *Journal of the National Medical Association, 99,* 1013–1023.
Stothard, S., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading & Writing, 4,* 245–256.
Swisher, L., Plante, E., & Lowell, S. (1994). Nonlinguistic deficits of children with language disorders complicate the interpretation of their nonverbal IQ scores. *Language, Speech, and Hearing Services in Schools, 25,* 235–240.
Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1275–1278.
Thordardottir, E., & Ellis Weismer, S. (2002). Verb argument structure weakness in specific language impairment in relation to age and utterance length. *Clinical Linguistics & Phonetics, 16,* 233–250.
Traxler, M., & Tooley, K. (2007). Lexical mediation and context effects in sentence processing. *Brain Research, 1146,* 59–72.
van der Lely, H. (1996). Specifically language impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua, 98,* 243–272.
van der Lely, H., & Harris, M. (1990). Comprehension of reversible sentences in specifically language impaired children. *Journal of Speech and Hearing Disorders, 55,* 101–117.
van der Lely, H., & Stollwerck, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition, 62,* 245–290.
Van Dyke, J. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 33, 407–430.

Van Dyke, J., & Lewis, R. (2003). Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalysed ambiguities. *Journal of Memory and Language, 49,* 285–413.

Van Dyke, J., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language, 55,* 157–166.

Van Dyke, J., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language, 65,* 247–263.

Vitevitch, M., & Luce, P. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36,* 481–487.

Wallace, G., & Hammil, D. (2000). *Comprehensive Receptive and Expressive Vocabulary Test–Second Edition.* Austin, TX: Pro-Ed.

West, S., Finch, J., & Curran, P. (1996). Structural equation models with nonnormal variables: Problems and remedies. In R. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56–75). Thousand Oaks, CA: Sage.

Westby, C. (1994). The effects of culture on genre, structure, and style of oral and written texts. In G. Wallach & K. Butler (Eds.), *Language learning disabilities in school-age children and adolescents: Some principles and applications* (pp. 180–218). New York, NY: Merrill.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32,* 25–64.

Appendix

Sample experimental sentences

Subject–Verb–Object
The hat had hugged the belt behind the very bright new sock.
The ring had moved the square behind the very bright cold bed.
The square had changed the bed under the very new dry key.
Subject Focus Center Embedded Relatives
The watch that had hugged the truck behind the kite was bright.
The train that had helped the knife under the square was cold.
The boot that had fixed the shoe behind the drum was new.
Passive
The train was watched by the bed behind the very cold cake.
The watch was bumped by the wheel near the very bright clock.
The key was changed by the chair behind the very bright square.
Object Focus Center Embedded Relatives
The box that the kite had splashed behind the shoe was dry.
The truck that the clock had pressed near the door was bright.
The chair that the bread had splashed under the square was new.