ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Sentence Prediction Deficits in Developmental Language Disorder Are a Product of Vocabulary Knowledge and Processing Abilities

Justin B. Kueser,[a] Claney Outzen,[b] Arielle Borovsky,[b] Patricia Deevy,[b] and Laurence B. Leonard[b]

[a] Boys Town National Research Hospital, Omaha, NE [b] Purdue University, West Lafayette, IN

ARTICLE INFO

Article History:
Received February 18, 2025
Revision received October 16, 2025
Accepted February 3, 2026

Editor-in-Chief: Julie A. Washington
Editor: Raúl Rojas

https://doi.org/10.1044/2026_JSLHR-25-00121

ABSTRACT

Purpose: Children with developmental language disorder (DLD) have difficulty making online predictions of language material following verbs (e.g., "The monkey eats a very delicious ... [banana]"). We explore the contributions of lexicosemantic knowledge deficits and online processing deficits by comparing performance across offline lexicosemantic and online processing tasks in sentences with higher versus lower speed and complexity.
Method: Participants included twenty-six 4- to 5-year-old children with DLD and 26 age-matched children with typical development (TD). In Experiment 1, participants' lexicosemantic knowledge about verb–patient associates was assessed (e.g., "What do babies usually wear? A bib or a necklace?") in an offline pointing task. Following the offline task, participants' online predictive processing for the same verb–patient associates was assessed using eye tracking (e.g., "The baby is wearing a very special [bib vs. necklace]"). In Experiment 2, the same tasks were completed with simpler sentences spoken at a slower rate with a reduced number of words.
Results: Across the two experiments, the quality of lexicosemantic knowledge impacted the quality of sentence prediction for children in both groups. In addition, children with DLD had poorer lexicosemantic knowledge compared to peers with TD. Yet even after accounting for item-level lexicosemantic knowledge, the children with DLD differed from their peers with TD in sentence prediction. Specifically, children with DLD showed similar sentence prediction to peers with TD in faster sentences, but in slower sentences, their predictions decayed over time.
Conclusions: Children with DLD have sentence prediction deficits due to a combination of lexicosemantic knowledge deficits and problems with working memory decay and sustained attention. Sentence prediction deficits in DLD arise from both lexicosemantic and processing factors.
Supplemental Material: https://doi.org/10.23641/asha.31999236

Sentence comprehension is supported by both online processes—incremental processing that occurs as language input unfolds—and offline lexicosemantic knowledge—a store of nouns, verbs, and other words and their meanings accrued over time (Ferreira & Ferreira, 2024; MacWhinney, 2015). For children with developmental language disorder (DLD), a disorder of language learning, understanding, and use, both online processing and offline knowledge may be impaired, but how they interact is currently unknown. Thus, the goal of this study was to examine how offline lexicosemantic knowledge interacts with online language processing in children with DLD in the context of sentence comprehension.

When hearing a sentence beginning with "The monkey eats a very delicious ... ," many listeners might predict

Correspondence to Justin B. Kueser: justin.kueser@boystown.org.
Disclosure: *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

that the following word is "banana." In fact, when people hear sentences in visual world paradigm (VWP) studies in which images of alternatives are presented on-screen and eye gaze is measured, they tend to look at the object that is the most likely sentence-final word even before it is heard (e.g., Altmann & Kamide, 1999). Children with typical language development (TD) develop the ability to predict sentence completions early; toddlers as young as age 24 months are able to predict completions to sentences following verbs (Mani & Huettig, 2012).

Compared to children with TD, children with DLD show decreased speed, accuracy, and stability during prediction of nouns following verbs in VWP designs (Andreu et al., 2013; Borovsky et al., 2013; Kueser et al., 2024; van Alphen et al., 2021). Preschool-age and school-age children with DLD (vs. TD) are slower to use verb-mediated prediction to anticipate sentence-final words in both transitive (Andreu et al., 2013) and passive sentences (Kueser et al., 2024). Differences related to DLD (vs. TD) in verb-mediated prediction appear to be long-lasting, at least into adolescence (Borovsky et al., 2013) and likely beyond. The degree of prediction of sentence-final words is related to children's ability to recognize words in isolation, which is slower in children with DLD compared to their peers with TD (van Alphen et al., 2021). The fact that processing speed differences are observed in word recognition in addition to incremental prediction in sentences suggests that factors underlying word recognition—factors such as vocabulary knowledge and online processing speed—may help explain sentence-prediction deficits in DLD.

Developmental models of sentence processing find that repeated poor predictions may have cascading negative effects on language development due to misallocated attention and lost opportunities for learning (Bannon et al., 2025; Dell & Chang, 2014). Given the links between language learning and prediction, understanding why children with DLD have poor sentence-prediction ability provides both opportunities for theoretical tests concerning the role of prediction in learning, as well as translational potential to identify therapeutic targets for intervention. However, at present, it is not clear why children with DLD have poorer sentence prediction than children with TD, posing a barrier to implementing interventions focused on improving this skill. There are two reasonable accounts that have not yet been distinguished in the literature. The first is that children with DLD have less/under-specified lexicosemantic knowledge compared to peers with TD. The second is that children with DLD have sufficient lexicosemantic knowledge but experience cognitive or processing challenges in using their lexicosemantic knowledge to support predictive processing. We consider each alternative below.

Hypothesis 1: Children With DLD Do Not Have the Lexicosemantic Knowledge Necessary to Make Robust Predictions

Predicting upcoming patients at the end of a sentence such as "The monkey eats a . . . " requires a sizeable network of lexicosemantic knowledge, including knowledge of the semantics of the verb's patient thematic role (e.g., that things that are eaten tend to be foods) and knowledge of the semantics of potential noun completions (e.g., that a banana is a food and a rock usually is not), associations that are tracked by children (Borovsky, 2022; Yazbec et al., 2019). Work with toddlers and older children with TD shows a positive relationship between indices of vocabulary quality and sentence prediction (Borovsky et al., 2012; Gambi et al., 2021; Mani & Huettig, 2012; Reuter et al., 2023). In typical development, toddlers begin noun and verb learning with word-to-world mapping, learning verbs that describe actions that are highly individual in the environment (e.g., "kiss"; Gleitman et al., 2005). Children then engage in structure-to-world mapping in which syntactic structure, including verb argument structure, are used to learn words (e.g., Naigles, 1990). As these word-learning tools develop, children establish conceptual links between noun and verb meanings, resulting in growing networks of lexicosemantic knowledge that connect noun and verb meanings (Kueser, Horvath, & Borovsky, 2023), potentially serving as semantic material for sentence-level prediction in later development.

However, learning the semantic constraints of verbs is a challenge, particularly for children with DLD. Learning verb meaning is complicated by the fact that the events they describe are inherently relational, transitory, and dynamic in nature, with many involving multiple participants (Childers & Tomasello, 2006; Gentner, 1982; Gentner & Boroditsky, 2001). In fact, based on an analysis of the semantic features of early-learned nouns and verbs, the semantics of verb thematic roles are substantially more complex than the semantics of nouns (Kueser & Borovsky, 2025). In general, children with DLD have poorer noun and verb vocabulary and semantic feature knowledge than children with TD (Alt et al., 2004; Kan & Windsor, 2010; McGregor et al., 2012). In addition, late talkers, many of whom go on to develop DLD, also have differences in verb semantics compared to toddlers with TD (Horvath et al., 2019, 2022). Work by Pamela Hadley and colleagues supports the idea that children with language difficulties have poorer knowledge of the semantics associated with verbs' arguments (Hadley, 2020). For example, late talkers have smaller verb vocabularies than typically developing children (Hadley et al., 2016). Diversity among the grammatical subjects produced by toddlers is positively associated with metrics of toddlers' grammatical development such as mean length of utterance

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

(MLU), suggesting that growth in the lexicosemantic knowledge of verbs' argument positions is linked to better language abilities (Hadley et al., 2018).

Thus, one possible explanation for the poorer sentence prediction ability in children with DLD is that they have less detailed or robust lexicosemantic knowledge, particularly of verbs. One study has examined how stimuli-specific lexicosemantic knowledge relates to prediction ability in DLD (Kueser et al., 2024). This study found that knowledge of the specific verb used in the sentence was positively correlated with sentence prediction performance in children with DLD, suggesting a tight linkage between lexicosemantic knowledge of words in a sentence and the robustness of sentence prediction.

However, children with DLD often have other language processing difficulties in addition to vocabulary deficits. In addition, as it is yet unclear how much vocabulary knowledge one needs in order to predict, it could be that children with DLD have enough lexicosemantic knowledge to get by but have difficulties with prediction for other reasons.

Hypothesis 2: Children With DLD Have Difficulty Using Lexicosemantic Knowledge to Make Predictions in the Moment

In addition to lexicosemantic knowledge, predicting upcoming words requires the efficient coordination of language and cognition (Huettig et al., 2022). First, listeners must note the general class of word to be predicted (e.g., noun) and activate the relevant semantic features of the predicate (e.g., <food> for *eat*). Second, listeners must activate (or otherwise prepare for processing) words that have these properties while keeping these predictions active in working memory as potentially other words are heard. For VWP studies, a third step is to match the preactivated words or semantic features to pictures in the visual scene.

Beyond difficulty with language, children with DLD show differences with other cognitive processes, including slower processing speed (Leonard et al., 2007; Miller et al., 2001, 2006; Montgomery & Windsor, 2007; Park et al., 2020; Zapparrata et al., 2023), vulnerability to interference (e.g., Gillam et al., 1995; Marton et al., 2014, 2016; Seiger-Gardner & Brooks, 2008; Seiger-Gardner & Schwartz, 2008), difficulty maintaining material in working memory over time (Bishop et al., 1996; Dollaghan & Campbell, 1998; Gillam et al., 1998; Leonard et al., 2007; McMurray et al., 2010; cf. Alt & Spaulding, 2011; Poll et al., 2014; Seiger-Gardner & Schwartz, 2008), and problems with inhibition (Im-Bolter et al., 2006; McMurray et al., 2010, 2014; Spaulding, 2010). Any or several of these cognitive processes shown to be impaired in children with DLD could affect prediction. Here, we focus on three broad factors likely to be relevant to sentence prediction: (a) quick processing speed and working memory robust to interference (Hypothesis 2A), (b) the maintenance of predictions in working memory across time and adequate sustained attention (Hypothesis 2B), and (c) the activation of relevant and inhibition of irrelevant semantic features (Hypothesis 2C).

Hypothesis 2A: Children With DLD Have Difficulty Making Predictions in Rapidly Unfolding Sentences Due to Slow Processing Speed and Working Memory Interference

Sentences presented at a rapid pace can pose problems in prediction for children with DLD for two reasons: First, linguistic material may occur too quickly for children with DLD to keep up with, and second, rapidly incoming linguistic material may overwrite or interfere with items already in working memory. Children with DLD have slower overall processing speed compared to peers with TD in a variety of tasks (Leonard et al., 2007; Miller et al., 2001, 2006; Montgomery & Windsor, 2007; Park et al., 2020; Zapparrata et al., 2023). One implication of this is that the sentence prediction studies showing deficits in online prediction could have presented stimuli at a rate that is too fast for individuals with DLD to process. In fact, in several studies, children with DLD demonstrated better language processing when material was presented at slower rates (Montgomery, 2004, 2005; Poll et al., 2014; Witherstone, 2024). These studies suggest that if sentences are slowed down, then children with DLD may have similarly robust online prediction compared to children with TD.

Children with DLD also have difficulties managing interference across words and trials in verbal tasks (e.g., Gillam et al., 1995; Marton et al., 2014, 2016; Seiger-Gardner & Brooks, 2008; Seiger-Gardner & Schwartz, 2008). In these studies, children with DLD show slowed or inaccurate responses when words or experimental trials share features in memory, which produces interference that can negatively affect working memory (Oberauer & Lewandowsky, 2014). In some studies that show differences between DLD and TD groups in predictive ability, additional words intervened between the verb and the word-to-be-predicted (e.g., Andreu et al., 2013; Kueser et al., 2024). Difficulty with online sentence prediction in these studies may have arisen from difficulties that children with DLD experience in maintaining a prediction in working memory in the face of interference from additional words; thus, if sentences are simplified, then children with DLD may have similar sentence prediction as children with TD.

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Hypothesis 2B: Children With DLD Have Difficulty Making Predictions in Slower Sentences Due to Memory Decay and Poor Sustained Attention

Hypothesis 2B is based on two pieces of contrary evidence that suggest that children with DLD may have more difficulty processing language material that takes longer to unfold in time. The first piece of evidence comes from studies suggesting that memory traces, particularly when they involve phonological material, decay more quickly in children with DLD compared to children with TD (Bishop et al., 1996; Dollaghan & Campbell, 1998; Gillam et al., 1998; Leonard et al., 2007; McMurray et al., 2010; cf. Alt & Spaulding, 2011; Poll et al., 2014; Seiger-Gardner & Schwartz, 2008). The second piece of evidence comes from work showing that children with DLD have difficulties with sustained attention for tasks involving auditory and visual information (Ebert & Kohnert, 2011; Finneran et al., 2009; Smolak et al., 2020). Together, these results suggest that children with DLD may have difficulty with slowly (vs. rapidly) presented language material because there are more opportunities for memory traces to fade and attention to wander.

Hypothesis 2C: Children With DLD Have Difficulty Making Fine-Grained Predictions Due to Problems With Online Activation and Inhibition of Semantic Features

Relative to children with TD, children with DLD may also have challenges with graded prediction, which is the weighted, simultaneous pre-activation of multiple competing alternatives that reflects differences in associative strengths between the context and the items-to-be-predicted (Carter et al., 2019; Gambi et al., 2021; Luke & Christianson, 2016; Staub, 2015; Szewczyk & Federmeier, 2022). For example, in a VWP study, looks toward a "banana" would likely exceed looks toward an "apple" when predicting the completion of the sentence "The monkey is eating a … ," despite the fact that monkeys could plausibly eat both apples and bananas.

There is evidence to suggest that both lexicosemantic knowledge and processing factors may make graded prediction less accurate and/or efficient for children with DLD compared to peers. One study that has examined aspects of graded prediction in children with DLD (Andreu et al., 2013) contrasted predictions of typical and atypical sentence completions in 5- to 8-year-old Spanish-speaking children with DLD, age-matched peers with TD, younger children with TD matched on MLU, and adults. They found evidence for better graded prediction in the adults and age-matched children compared to the group with DLD, but similar graded prediction to younger children with TD matched on MLU, which is consistent with an account that vocabulary deficits and not processing deficits were primary in affecting graded prediction in children with DLD.

Other evidence suggests that processing factors such as inhibition of alternative meanings may underlie graded prediction deficits in DLD. Across a wide variety of tasks, children with DLD demonstrate difficulties with inhibition (Im-Bolter et al., 2006; Spaulding, 2010), including in lexical recognition tasks (McMurray et al., 2010, 2014). Thus, problems with inhibiting irrelevant semantic information may cause these children to demonstrate less sensitive sentence prediction, particularly when alternatives are close in associative strength with the verb or when they share semantic features.

Current Study

The hypotheses above make clear predictions about processes that rely on stored semantic knowledge—which should be evident even in offline language tasks where timing pressure is at a minimum—and among processing-related skills that emerge during real-time comprehension—that is, during online tasks. Therefore, to distinguish between these hypotheses, we separately examine offline lexicosemantic knowledge and online predictive processing under different demands in 4- to 5-year-old children with and without DLD. In the offline lexicosemantic knowledge task, we assess the specificity of children's knowledge of patients associated with early-learned agents and verbs. Children choose the patient most associated with a verb among nouns that have high, medium, or low association with the verb (e.g., "What does a horse usually eat, high: hay or low: bacon? What does a baby usually wear, high: a diaper or medium: a coat?"). These nouns are arrayed in pairs that make distinguishing the correct option easier (high vs. low associates) or harder (medium vs. low associates). In the online predictive processing task, we use eye tracking to assess how well children predict the noun following the agent and verb, using the same nouns as in the offline lexicosemantic knowledge task (e.g., "The horse is eating the __; The baby is wearing the __"). By specifically examining online prediction trials in which children show correct knowledge in the offline lexicosemantic knowledge task, we will be able to make inferences about the effect of lexicosemantic knowledge separate from the effect of children's processing abilities. Finally, to further test the hypothesized mechanisms involved in predictive processing in DLD, the sentences vary in their duration and in the specific stimuli used across two experiments. Experiment 1 uses more complex sentences with a faster rate while Experiment 2 uses simpler sentences with a slower rate.

Each hypothesis would be consistent with unique patterns of performance across tasks and experiments (see Table 1). Hypothesis 1 would be supported by a pattern

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 1. Outline of hypotheses and results.

Hypothesis	Prediction by task			Support from experimental data
	Experiments 1 and 2: offline knowledge	Experiment 1: online prediction in faster, more complex sentences	Experiment 2: online prediction in slower, less complex sentences	
1. Poor lexicosemantic knowledge	TD > DLD	For sentences with good offline knowledge, TD = DLD	For sentences with good offline knowledge, TD = DLD	Strong support from offline knowledge
2A. Slow processing and working memory interference	TD = DLD	For sentences with good offline knowledge, TD > DLD for all trials	For sentences with good offline knowledge, TD = DLD for all trials	No support from offline knowledge or online prediction
2B. Memory decay and poor sustained attention	TD = DLD	For sentences with good offline knowledge, TD = DLD for all trials	For sentences with good offline knowledge, TD > DLD for all trials	Strong support from online prediction
2C. Problems with online activation/inhibition	TD = DLD	For sentences with good offline knowledge, TD > DLD across different trial types (H/L vs. H/M vs. M/L)	For sentences with good offline knowledge, TD > DLD across different trial types (H/L vs. H/M vs. M/L)	Partial support from online prediction

Note. TD = typical language development; DLD = developmental language disorder; H/L = high–low; H/M = high–medium; M/L = medium–low.

of less accurate offline lexicosemantic performance in children with DLD compared to peers with TD, as well as equivalent performance across groups in the online tasks for items that were answered correctly in the offline lexicosemantic knowledge task. Alternatively, based on Hypothesis 2, we would predict that even for items answered correctly in the offline lexicosemantic knowledge task, children with DLD will show poorer online predictive processing than children with TD. Hypothesis 2A would lead to a prediction that poorer online predictive processing will be more evident in faster sentences compared to slower sentences for children with DLD specifically. Hypothesis 2B would be consistent with a pattern where children with DLD would show relatively good performance (compared to peers) in fast sentences where features in memory have not had (as much) time to decay but worse performance in slower sentences where features in memory may be lost. Hypothesis 2C would predict that online predictive processing will vary across different trial types (e.g., high vs. low associates compared to high vs. medium associates compared to medium vs. low associates: "The horse is eating the hay/bacon vs. hay/orange vs. orange/bacon"). If online predictive processing is worse in trials with larger differences in associative strength (e.g., high vs. low associates), then that would suggest a failure to appropriately weight clear predictive information. In contrast, if processing is worse in trials with smaller differences in associative strength (e.g., high vs. medium or medium vs. low associates), then that would suggest a failure to detect and act on the small difference. Hypotheses 2A–2C are agnostic with respect to offline knowledge, but given that they are focused on issues with online mechanisms, they seem most consistent with a prediction that offline knowledge is roughly equivalent in both groups.

Experiment 1: Prediction in Fast Sentences

Method

This study was conducted in accordance with the institutional review board at Purdue University (Institutional Review Board Protocol 1904021980). Participants' caregivers provided written consent, and participants provided verbal assent before each session.

Participants

Participants were 4- to 5-year-old children—14 children with DLD (six girls, eight boys) and 14 age-matched children with TD (seven girls, seven boys; see Table 2 for demographic and test score data). The study was conducted in American English. Participants were users of Midwestern American English typical of Indiana (N = 28). Participants were White (n = 23) and multiple races (n = 5), with two also being Hispanic. The average number of years of maternal education was matched between groups.

The children with DLD were required to score below 87 on the Structured Photographic Expressive Language Test–Preschool 2 (SPELT-P 2; Dawson et al., 2005); this is the empirically derived cutoff score for high sensitivity and specificity for this test (Greenslade et al., 2009). The children with TD were required to score above

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 2. Demographic and test score data across experiments.

Experiment	Measure	DLD		TD				
		M	*SD*	*M*	*SD*	*d*	*t*	*p*
Experiment 1	Age (months)	58.00	6.67	58.40	7.02	–0.05	–0.14	.891
	Maternal education (years)	16.20	2.67	16.60	1.65	–0.16	–0.43	.674
	SPELT-P 2 (SS)	71.90	12.10	120.00	5.33	–5.15	–13.62	< .001 ***
	KABC-2 (SS)	101.00	17.80	119.00	10.30	–1.27	–3.36	.003 **
	PPVT-5 (SS)	95.40	13.90	124.00	12.50	–2.20	–5.82	< .001 ***
	EVT-3 (SS)	93.90	8.98	119.00	14.80	–2.06	–5.45	< .001 ***
Experiment 2	Age (months)	57.33	5.60	59.33	4.85	–0.38	–0.94	.360
	Maternal education (years)	16.58	2.78	16.96	1.74	–0.16	–0.40	.696
	SPELT-P 2 (SS)	76.58	10.49	116.33	7.28	–4.40	–10.78	< .001 ***
	KABC-2 (SS)	105.50	5.23	112.58	12.97	–0.72	–1.75	.100
	PPVT-5 (SS)	97.92	12.24	117.67	9.98	–1.77	–4.33	< .001 ***
	EVT-3 (SS)	97.58	10.90	115.58	17.44	–1.24	–3.03	.007 **

Note. DLD = developmental language disorder; TD = typical language development; SPELT-P 2 = Structured Photographic Expressive Language Test–Preschool 2; SS = standard score; KABC-2 = Kaufman Assessment Battery for Children–Second Edition; PPVT-5 = Peabody Picture Vocabulary Test–Fifth Edition; EVT-3 = Expressive Vocabulary Test–Third Edition.

$^{**}p < .01$. $^{***}p < .001$.

87 on the SPELT-P 2 and have no parental concerns about language development. The children with DLD were also required to have minimal or no symptoms of autism spectrum disorder on the Childhood Autism Rating Scale–Second Edition (Schopler et al., 2010).

All children were required to pass a pure-tone hearing screening at 1, 2, and 4 kHz; have no history of neurological disorders; and achieve a standard score above 75 on the Kaufman Assessment Battery for Children–Second Edition (Kaufman & Kaufman, 2018), a measure of nonverbal intelligence. All children were additionally administered the Peabody Picture Vocabulary Test–Fifth Edition (Dunn, 2019) and the Expressive Vocabulary Test–Third Edition (Williams, 2019) as supplemental measures of receptive and expressive vocabulary knowledge.

Stimuli

We selected four early-learned verbs (*eat*, *hold*, *ride*, and *wear*), and for each verb, we selected three unique agents. For each agent–verb combination, we selected a large number of candidate patients and elicited ratings on a 5-point scale from *very unlikely* to *very likely* from a minimum of 19 adult English-speaking participants from Prolific about two properties of the stimuli. First, we asked participants about the likelihood that the agent would interact with each patient (agent–patient associations). Second, we asked for the likelihood that the agent would perform the action on each patient (agent–verb–patient associations). Based on these ratings, we selected six patients for each agent–verb combination, categorized as having high, medium, or low agent–patient and agent–verb–patient associations. High-, medium-, and low-associated patients had significantly different ratings from one another (see Table 3), agent–patient association scores: $p < .001$, agent–verb–patient association scores, $p < .001$. For each verb, there were 12 potential pairs of patients that differed in their association category (i.e., four high–low [HL] pairs, four high–medium [HM] pairs, and four medium–low [ML] pairs). The verbs, agents, and patients along with their agent–patient and agent–verb–patient association scores are provided on the Open Science Framework

Table 3. Agent–patient and agent–verb–patient association scores as rated by adults across experiments.

Experiment	Condition	Agent–patient association score			Agent–verb–patient association score		
		Mdn	*M*	*SE*	*Mdn*	*M*	*SE*
Experiment 1	High	5.00	4.67	0.95	5.00	4.76	0.97
	Medium	3.00	3.01	0.61	3.00	3.24	0.66
	Low	1.00	2.00	0.41	1.00	1.48	0.30
Experiment 2	High	5.00	4.67	0.95	5.00	4.74	0.97
	Medium	3.00	3.00	0.61	3.25	3.20	0.65
	Low	1.00	2.00	0.41	1.00	1.54	0.32

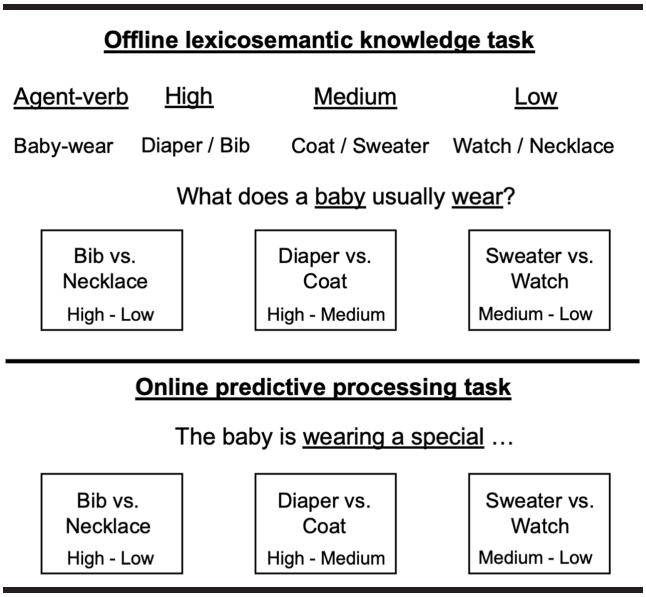
Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

(OSF; https://osf.io/cmesb/). Last, visual stimuli were created for each patient using 400 × 400 pixel color photographs (see Figure 1, e.g., stimuli pairs).

Offline Lexicosemantic Knowledge Task

The offline lexicosemantic knowledge task was conducted on the first day of the study. The purpose of this task was to establish children's knowledge of agent–verb–patient associations. Trials consisted of two parts (see Figure 1 for outline of experimental design). First, for each agent–verb–patient pair presented to a child (e.g., HL pair for horse/eat: hay vs. bacon), each patient was explicitly named as its picture was displayed on a computer screen. Second, the children were asked to point to a picture in response to the question "Which one does a [agent] usually [verb]?" (e.g., "Which one does a horse usually eat?"). The children's responses were recorded by an examiner with a button box. Before the test trials, the children were presented with two initial training trials (*eat–girl–oatmeal* vs. *paper* and *carry–person–box* vs. *mountain*) and feedback was provided; if their answers were incorrect (e.g., "people carry mountains"), they were prompted to try again until they provided the correct answer. Each child saw half of the potential patient pairs per verb (i.e., one HL, one HM, and one ML pairs) such that each patient was seen by a child in only one pair. Four counterbalance orders were created by varying the specific patient pairs seen for each verb and by varying the order of the verbs. There were 36 trials for this task (4 verbs × 3 agents/verb × 3 patient pairs/verb).

Figure 1. Example stimuli pairs and outline of experimental design.



Binary accuracy data (correct vs. incorrect) were analyzed using Bayesian mixed-effects logistic regression. A Bayesian statistical approach was selected primarily because it accommodated modeling the eye tracking data from the online task using a beta distribution with a complex random-effects structure that could not be successfully fit using frequentist statistical packages. Other advantages of a Bayesian approach are reviewed in Kruschke and Liddell (2018). We included random intercepts for participant and item embedded in agent and random slopes for group on the agent/item intercept and condition on the participant intercept. Independent variables were group (DLD vs. TD), condition (HL vs. HM vs. ML), and their interaction. The model equation in Wilkinson notation is $Accuracy \sim Condition \times Group + (Condition \mid Participant) + (Group \mid Agent/Item)$. Condition and group were sum-to-zero coded. As condition is a three-level variable, the regression coefficients reflect sum-to-zero coding (i.e., Condition1 = [HL = 1, HM = 0, ML = −1]; Condition2 = [HL = 0, HM = 1, ML = −1]) and are only minimally informative. Therefore, we will report pairwise differences between conditions regardless of the values in the regression coefficient tables. There were 1,008 observations for this model. As priors on the fixed effects, we used a weakly informative normal distribution centered on 0 with an *SD* of 0.5. We collected 2,000 warmup samples and 10,000 post-warmup samples for the model so that we would have adequate precision for 95% CIs.

To contextualize the strength of effects, we used credible intervals (CIs), p_d values ("p direction"), and p_{ROPE} values ("proportion of posterior samples in the region of practical equivalence"). All of these metrics help to determine the likelihood that effects are present or absent and to quantify their strength. CIs are similar to frequentist confidence intervals but have a simpler interpretation—at a given probability level, CIs represent the probability that the true parameter value is within the given range. We report 95% highest density CIs, meaning that there is a 95% likelihood that the true parameter value is within the interval. p_d Values are the proportion of posterior samples above or below zero for point estimates above or below zero, respectively. In our tables, we star p_d values greater than .95 to focus on their interpretation, though, in fitting with our Bayesian approach, effects with p_d values below .95 should not be considered "insignificant" or absent but rather as being weaker or having less evidence in support of them. We characterize comparisons with p_d values above .95 as having strong evidence, values from .90 to .95 as having weak evidence, and values below .90 as having little evidence. We also calculate p_{ROPE} for all post hoc comparisons. The region of practical equivalence (ROPE) is defined as the range of effect sizes considered to be negligible; p_{ROPE} is the

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

proportion of posterior samples within the negligible effect size range. Thus, when p_{ROPE} is high, there can be high confidence that the effect is at or near zero, and when p_{ROPE} is low, there can be high confidence that the effect is nonzero and meaningful. For the models in this article, effects are measured on the logit or log odds scale, and so we use a range of −0.055 to 0.055 for the ROPE (Cohen, 1988; Sánchez-Meca et al., 2003). The statistical analysis code is available at https://osf.io/cmesb/. We used *R* Version 4.4.2 (R Development Core Team, 2008), *emmeans* Version 1.10.5 (Lenth, 2016), *tidybayes* Version 3.0.7 (Kay, 2021), and bayestestR Version 0.15.0 (Makowski et al., 2019) to perform post hoc testing and *brms* Version 2.22.0 (Bürkner, 2018) to fit the statistical models.

Online Predictive Processing Task

The online predictive processing task was administered at least 1 week after the offline lexicosemantic knowledge task. The purpose of this task was to measure the influence of offline knowledge on online sentence processing and to assess group differences in online processing in sentences with good offline knowledge. To ensure a tight linkage with the offline knowledge task, we used the same agent–verb–patient pairs. The same visual stimuli were used to represent each patient. Trials consisted of two parts, a preview phase and a test phase. The preview phase displayed the two images on-screen for 2,000 ms. Following the preview period, a 100 × 100 pixel animated central fixation appeared on the screen between the two images. When children looked at this video for 200 ms, it disappeared, and the test phase of the trial began. In the test phase, the two images continued to be visible. An audio stimulus consisting of a sentence with the following form played at 70 dB: "The [agent] is [verb]+ing a [adjective phrase] [patient]." Informative portions of the sentences were standardized in duration (see Table 4). Adjective phrases were included to increase the length and complexity of the sentence while being semantically neutral with respect to the patients (e.g., "really good," "very special"). There were 36 trials for this task (4 verbs × 3 agents/verb × 3 patient pairs/verb).

Looks toward the target-versus-distractor patient were recorded during the preview and test phases at 500 Hz using an EyeLink 1000 Plus eye tracking system, and a 24-in. screen was positioned 580–620 mm away from the children. The system was calibrated using a 5-point routine, and a central fixation image was presented before every trial to calibrate the machine in case of drift or movement. An experimenter sat next to the child and provided noncontingent praise. Trials were excluded from analysis if there was more than 80% track loss (e.g., eye blinks or head turns away from screen) across the sentence. This criterion resulted in the exclusion of 41 out of 1,008 trials (4% of trials).

Table 4. Durations of parts of sentences across experiments.

Sentence part	Experiment 1	Experiment 2
First noun (agent) + auxiliary	1,196	2,714
Verb	681	1,384
Determiner (+ adjective)	1,211	1,006
Final noun (patient)	912	1,258
Total duration	4,000	6,362
Total postverb predictive	1,892	2,390
Total prefinal noun predictive	3,088	5,104

Note. Durations are in milliseconds. Adjective was present in Experiment 1 only.

The looking data were prepared for analysis by binning the 500-Hz raw samples into intervals of 50 ms, producing a proportion of target and distractor looking every 50 ms during the sentence, following prior work (e.g., Borovsky et al., 2013). Next, the proportion of target looking was divided by the sum of the proportion of target and distractor looking to reflect bias toward the target image in predictive processing; we refer to this measure as "target proportion." Last, average target proportion was calculated for each sentence in two windows: the 2,000-ms preview window prior to the appearance of the central fixation video and the 1,892-ms test window between the onset of the verb and the onset of the patient (i.e., "[verb]+ing a [adjective phrase]").

We analyzed the target proportion data in two models. Both models examined the change in target proportion from the preview to the test window to control for potential biases to individual pictures; evidence of predictive processing would emerge if target proportion in the test window exceeded that in the preview window. Both models used weakly informative normal distributions centered on 0, with an *SD* of 0.5 for the fixed effects and 10,000 post-warmup samples.

The first model ("Influence of offline knowledge") was focused on determining whether offline lexicosemantic knowledge was related to online predictive processing, a key assumption of our argument. This model was a Bayesian mixed-effects beta model with 967 trials included. The dependent variable was target proportion. The independent variables were offline accuracy (correct vs. incorrect on the offline lexicosemantic knowledge task for that item for that participant), group (DLD vs. TD), window (preview vs. test), and their interactions. Offline accuracy, group, and window were sum-to-zero coded. Random intercepts were participant and item embedded in agent, and random slopes included offline accuracy and window for the participant intercept and offline accuracy, group, and window for the agent/item intercept. The model

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

equation in the Wilkinson notation was *Target_proportion* ~ *Offline_accuracy* × *Group* × *Window* + (*Offline_accuracy* + *Window* | *Participant*) + (*Offline_accuracy* + *Group* + *Window* | *Agent/Item*).

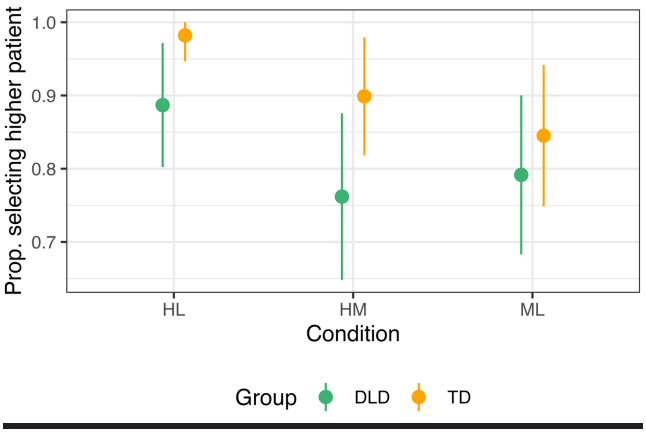
The second model ("Online processing with accurate knowledge") focused on determining whether online processing differed in the two groups for items for which participants had correct responses in the offline lexicosemantic knowledge task (N = 839 trials). This model was also a Bayesian mixed-effects beta model with target proportion as the dependent variable. The independent variables were condition (HL vs. HM vs. ML), group (TD vs. DLD), window (preview vs. test), and their interactions. Condition, group, and window were sum-to-zero coded. Random intercepts were participant and item nested in agent, and random slopes included condition and window for the participant intercept and group and window for the agent/item intercept. The model equation in the Wilkinson notation was *Target_proportion* ~ *Condition* × *Group* × *Window* + (*Condition* + *Window* | *Participant*) + (*Group* + *Window* | *Agent/Item*).

Results

Offline Lexicosemantic Knowledge Task

Figure 2 shows the proportion of trials within each condition in which the children selected the higher associated patient in the offline knowledge task. The mixed-effects logistic regression model showed strong evidence for differences in accuracy across conditions and between groups (see Table 5 for full model results). Specifically, across groups, accuracy in the HL condition, EMM (estimated marginal median) = .97, 95% CI [.94, .99], was higher than in the HM condition, EMM = .92, 95% CI [.85, .97], p_d = .996, p_{ROPE} = .004. (Recall that p_d denotes the proportion of posterior samples that are in the direction of the effect—and so larger proportions reflect stronger, more reliable effects—while p_{ROPE} denotes the proportion of posterior samples that are in a range corresponding to negligible effect sizes—and so smaller proportions reflect stronger, more reliable effects.) Compared to the HL condition, accuracy in the ML condition was also lower, EMM = .90, 95% CI [.82, .96], p_d = .998, p_{ROPE} = .001, but there was less evidence that the HM and ML conditions differed, p_d = .743, p_{ROPE} = .082. Across conditions, accuracy for the group with TD, EMM = .97, 95% CI [.93, .99], was higher than for the group with DLD, EMM = .88, 95% CI [.82, .94], p_d = .997, p_{ROPE} = .002. There was strong evidence for group differences in the HL and HM conditions but no evidence for group differences in the ML condition (see Table 6 for group comparisons by condition).

Figure 2. Accuracy on offline lexicosemantic knowledge task in Experiment 1. Error bars are 95% confidence intervals. Raw participant-level data are presented in Supplemental Figure S1 (see Supplemental Material S1). Prop. = proportion; HL = high–low; HM = high–medium; ML = medium–low; DLD = developmental language disorder; TD = typical language development.



Online Predictive Processing Task

Influence of offline knowledge. Figure 3 shows the proportion of target looking for each group in the preview and predictive windows for sentences with correct and incorrect answers on the offline lexicosemantic knowledge task. The mixed-effects beta regression model showed strong evidence for the influence of offline accuracy on online predictive processing across groups (see Table 7 for full model results). Specifically, there was strong evidence for an interaction between window and offline accuracy such that, across groups, in the predictive window, sentences with correct offline accuracy had greater proportions of target looking compared to sentences with incorrect offline accuracy, OR (odds ratio) = 2.08, 95% CI [1.38, 2.97], p_d = 1.00, p_{ROPE} < .001, but there was less evidence for such a difference in the preview window, OR = 1.20, 95% CI [0.79, 1.70], p_d = .839, p_{ROPE} = .144 (see Table 8 for effects broken down by group).

Online processing with accurate knowledge. Figure 4 shows the time course of target looking in sentences with accurate offline knowledge in the preview and predictive windows. Visual inspection suggests that both groups showed prediction of the final patient in the predictive window, though the group with DLD appeared to demonstrate a less smooth upward trajectory compared to the group with TD. Table 9 shows the results of the mixed-effects beta regression model examining average target proportion across groups, conditions, and windows.

Figure 5 shows the model predictions for average target proportion across windows, groups, and conditions. There was strong evidence for an effect of window but no evidence for an effect of condition or group. Specifically, across groups and conditions, target looking was greater

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 5. Model results for offline lexicosemantic knowledge task across experiments.

Experiment	Term	Estimate	Error	95% CI	p_d	p_{ROPE}
Experiment 1	Intercept	2.67	0.33	[2.05, 3.34]	1*	0
	Condition1	0.77	0.23	[0.32, 1.23]	1*	< .001
	Condition2	–0.25	0.23	[–0.69, 0.21]	.860	.104
	Group	0.67	0.23	[0.21, 1.12]	.997*	.004
	Group × Condition1	0.26	0.20	[–0.14, 0.65]	.901*	.097
	Group × Condition2	0.07	0.20	[–0.32, 0.45]	.634	.212
Experiment 2	Intercept	2.94	0.33	[2.35, 3.62]	1*	0
	Condition1	0.59	0.27	[0.06, 1.13]	.986*	.016
	Condition2	–0.01	0.26	[–0.51, 0.51]	.506	.173
	Group	0.36	0.22	[–0.06, 0.82]	.954*	.051
	Group × Condition1	–0.13	0.20	[–0.53, 0.28]	.734	.176
	Group × Condition2	0.16	0.18	[–0.20, 0.52]	.811	.166

Note. Contrast coding was Condition1 = [HL = 1, HM = 0, ML = –1]; Condition2 = [HL = 0, HM = 1, ML = –1]; Group = [TD = 1, DLD = –1]. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; HL = high–low; HM = high–medium; ML = medium–low.

*p_d > .95. *p_d > .90.

in the predictive window, EMM = .62, 95% CI [.55, .69], than in the preview window, EMM = .49, 95% CI [.44, .54], p_d = .999, p_{ROPE} = .002. Although visual inspection of Figure 5 suggests small differences between conditions and groups, examination of target proportion across groups and conditions reveals largely similar gains in target proportion from preview to predictive windows (see Table 10 for target proportion broken down by group and condition). Table 11 shows differences between conditions in target looking for each group; the pattern of comparisons indicates that there was only weak evidence for differences in target proportion between the HL and HM conditions and between the HL and ML conditions for the group with TD but no evidence for the other conditions or for the group with DLD. Table 12 shows the three-way interaction between window, group, and condition, showing group differences in predictive looking by condition. It reveals that there was no evidence for group differences in online prediction across conditions. Thus, overall, there was strong evidence that both groups showed online prediction in all conditions, only weak evidence for condition differences within groups, and no evidence for group differences in online prediction.

Interim Discussion

Experiment 1 had three clear results. First, the lexicosemantic knowledge underlying online sentence prediction was less robust in children with DLD compared to children with TD. Second, lexicosemantic knowledge measured using an offline task meaningfully predicted sentence prediction in an online eye tracking task in both groups. Third, after removing trials that participants incorrectly answered in the offline lexicosemantic knowledge task, children with DLD and peers with TD had highly similar sentence-prediction abilities. While Experiment 1 provides initial evidence for Hypothesis 1 (lexicosemantic deficits) and against Hypothesis 2 (processing factors), the critical evidence for or against Hypothesis 2 will come from comparisons between Experiment 1 and Experiment 2, where sentences are longer in duration and spoken at a slower rate.

Table 6. Group comparisons for offline lexicosemantic knowledge task across experiments.

Experiment	Condition	$OR_{TD/DLD}$	95% CI	p_d	p_{ROPE}
Experiment 1	HL	6.29	[0.79, 19.10]	.997*	.001
	HM	4.29	[0.67, 11.70]	.992*	.004
	ML	1.96	[0.33, 5.30]	.883	.038
Experiment 2	HL	1.58	[0.22, 4.84]	.767	.052
	HM	2.76	[0.51, 7.57]	.965*	.015
	ML	1.89	[0.51, 4.74]	.905*	.039

Note. Odds ratio (*OR*) is the odds of correct response by the typical language development (TD) group over the odds of correct response by the developmental language disorder (DLD) group. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; HL = high–low; HM = high–medium; ML = medium–low.

*p_d > .95. *p_d > .90.

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Experiment 2: Prediction in Slower Sentences

Method

Participants

New participants were recruited for Experiment 2. Inclusionary and exclusionary criteria were identical to Experiment 1. There were 12 children with DLD (eight girls, four boys) and 12 age-matched children with TD (seven girls, five boys; see Table 2 for demographic and test score data). Participants were users of Midwestern American English typical of Indiana. Participants were White (n = 20), Black (n = 2), Asian (n = 1), and multiple races (n = 1), with one also being Hispanic. The average number of years of maternal education was matched between groups.

Stimuli

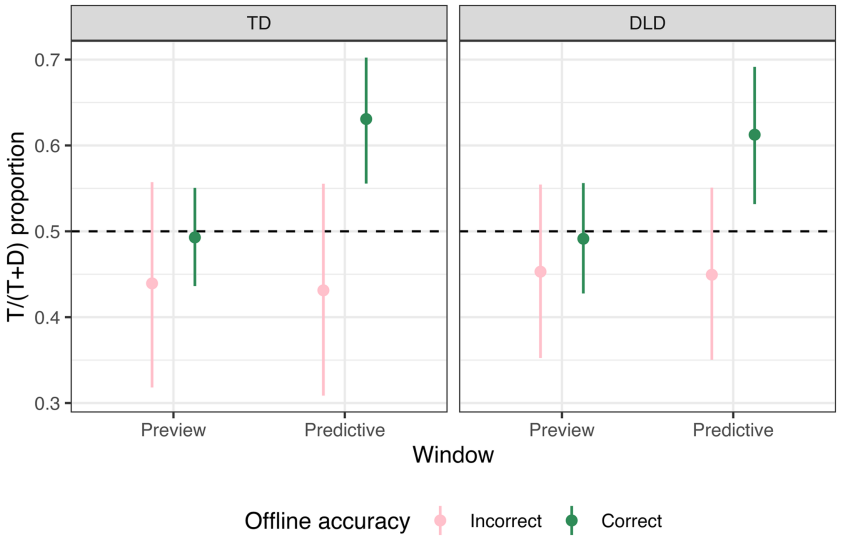
New stimuli were selected for Experiment 2 to demonstrate generalizability. Six verbs, two agents per verb, and six patients per agent were selected using the same stimuli selection procedure as for Experiment 2. High-, medium-, and low-associated patients had significantly different ratings from one another (see Table 3), agent–patient association scores: $p < .001$, agent–verb–patient association scores, $p < .001$. HL, HM, and ML patient pairs were chosen just as in Experiment 1. A list of the verbs, agents, and patients is provided on OSF (https://osf.io/cmesb/).

Offline and Online Tasks

The offline lexicosemantic knowledge task was identical to Experiment 1. There were 36 trials for this task (6 verbs × 2 agents/verb × 3 patient pairs/agent). The online predictive processing task was modified to decrease the rate and complexity of the sentences. As such, sentences took the form "The [agent] is [verb]+ing a [patient]," eliminating the adjective phrase, and they were recorded at a slower speaking rate. An additional change to the online predictive processing task was the inclusion of neutral sentences to better measure nonpredictive looking in response to the patient pairs. These sentences took the form "Someone is [neutral verb]+ing something." Neutral verbs were "look at" and "see." There was a neutral sentence included for each of the patient pairs. As in Experiment 1, the informative portions of the selective and neutral sentences were standardized in duration (see Table 4). Compared to Experiment 1, there were an additional 498 ms in the postverb predictive window and an additional 2,016 ms from the onset of the first noun to the final noun. This corresponds to an increase in duration of 26% for the postverb predictive window and 65% for all windows before the final noun relative to Experiment 1; this is similar to the difference between the slow and fast speech rate conditions in Montgomery (2005). There were 72 trials for this task (6 verbs × 2 agents/verb × 3 patient pairs/agent × 2 sentence types). The 36 selective trials were seen first, followed by the 36 neutral trials.

The data analysis procedure was identical to Experiment 1 with three exceptions. First, in the statistical

Figure 3. Model-predicted proportion of target looking for sentences with incorrect versus correct offline knowledge in Experiment 1. Raw participant-level data are presented in Supplemental Figure S2 (see Supplemental Material S1). Error bars are 95% confidence intervals. The horizontal dashed line indicates 50% target and 50% distractor looking. T / (T + D) = Target proportion / (Target proportion + Distractor proportion). TD = typical language development; DLD = developmental language disorder.



Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 7. Model results for influence of offline knowledge on online predictive processing across experiments.

Experiment	Term	Estimate	Error	95% CI	p_d	p_{ROPE}
Experiment 1	Intercept	0.00	0.09	[–0.18, 0.19]	.506	.458
	Offline accuracy	0.23	0.08	[0.06, 0.40]	.995*	.018
	Group	–0.01	0.07	[–0.14, 0.13]	.536	.579
	Window	0.13	0.07	[–0.02, 0.27]	.959*	.149
	Offline Accuracy × Group	0.03	0.07	[–0.11, 0.16]	.648	.553
	Offline Accuracy × Window	0.14	0.05	[0.05, –0.23]	.998*	.036
	Group × Window	0.01	0.06	[–0.11, 0.12]	.544	.645
	Offline Accuracy × Group × Window	0.01	0.04	[–0.08, 0.10]	.595	.767
Experiment 2	Intercept	0.23	0.07	[0.10, 0.36]	1*	.005
	Offline accuracy	0.04	0.06	[–0.07, 0.16]	.774	.524
	Group	0.04	0.06	[–0.08, 0.15]	.735	.560
	Sentence type	–0.03	0.06	[–0.15, 0.10]	.679	.579
	Offline Accuracy × Group	0.06	0.06	[–0.06, 0.17]	.869	.425
	Offline Accuracy × Sentence Type	0.09	0.05	[0.00, 0.19]	.970*	.218
	Group × Sentence Type	0.07	0.06	[–0.06, 0.19]	.861	.402
	Offline Accuracy × Group × Sentence Type	0.02	0.05	[–0.07, 0.12]	.669	.694

Note. Contrast coding was Offline accuracy = [Correct = 1, Incorrect = –1]; Group = [TD = 1, DLD = –1]; Window = [Predictive = 1, Preview = –1]; Sentence type = [Selective = 1, Neutral = –1]. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; TD = typical language development; DLD = developmental language disorder.

*$p_d > .95$.

models, the independent variable *window* was replaced with the variable *sentence type*. Second, the random intercept of *item* nested in *agent* was modified to include only *item* because neutral sentences only had one agent (i.e., "somebody"). Third, the number of trials excluded due to track loss and the procedure for handling excluded trials differed. For Experiment 2, 10 trials were removed due to track loss. To be included in the statistical models, we required that a participant contributed data for both the selective and neutral sentences for a patient pair. Thus, 10 additional trials where a participant contributed data for only the selective or neutral sentence were removed. These excluded trials represented 1.2% of the full data set (1,728 trials). For the offline lexicosemantic knowledge task, there were 864 trials included in the data analysis; for the online predictive processing task, there were 1,708 trials for the influence of the offline knowledge model and 1,474 trials for the online processing with accurate knowledge model.

Last, we ran two additional analyses, presented in Supplemental Material S1. The first supplemental analysis was focused on differences between prediction before the verb (i.e., from the agent) and prediction after the verb in order to explore the potential effect of decay of predictions in working memory across the sentence; results are referenced in the General Discussion section. The second supplemental analysis addresses the fact that the magnitude of difference between the groups with TD and DLD in their language test scores differs across experiments; this difference potentially complicates the comparisons we can make across experiments. As such, the analysis examines how continuous language scores (rather than group

Table 8. Incorrect versus correct offline accuracy comparisons for influence of offline knowledge on online predictive processing across experiments.

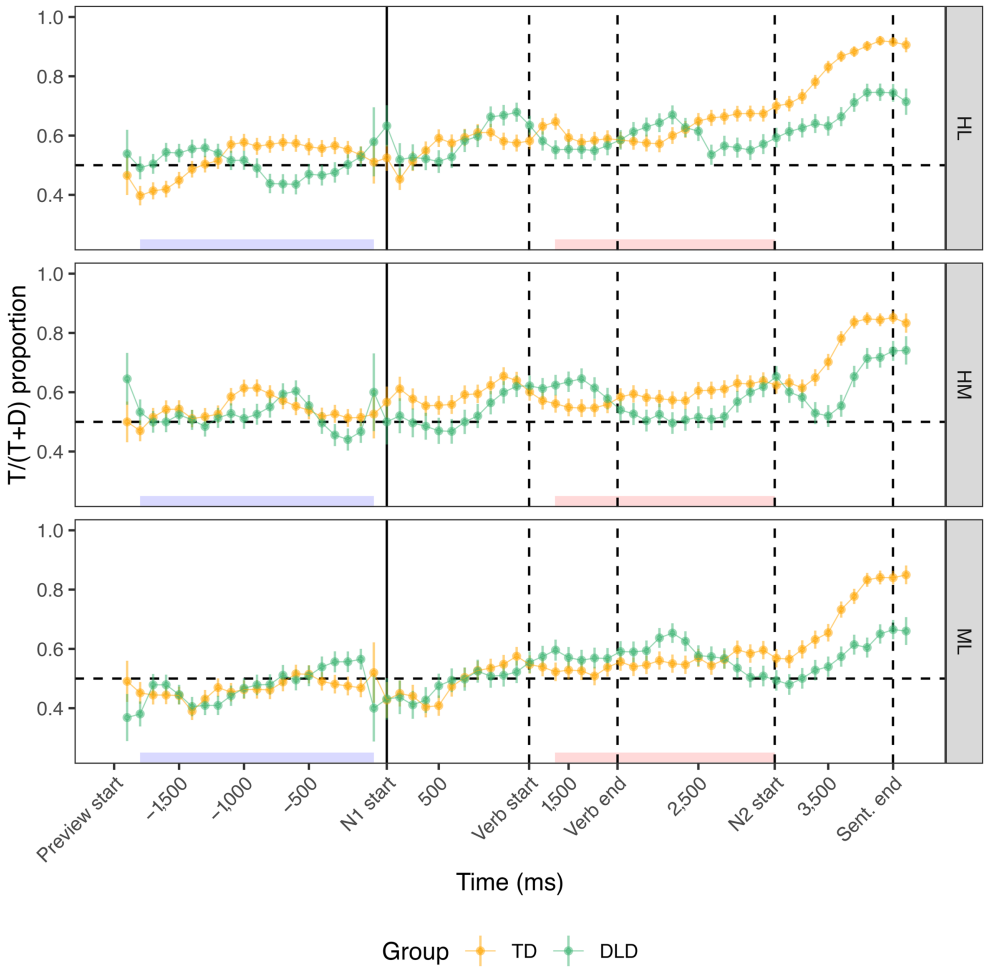
	Experiment 1					Experiment 2				
Group	Window	$OR_{Correct/Incorrect}$	95% CI	p_d	p_{ROPE}	Sentence type	$OR_{Correct/Incorrect}$	95% CI	p_d	p_{ROPE}
TD	Preview	1.24	[0.66, 2.00]	.788	.116	Neutral	0.98	[0.60, 1.46]	.532	.201
	Predictive	2.24	[1.18, 3.68]	.998*	.003	Selective	1.56	[0.92, 2.32]	.974*	.030
DLD	Preview	1.17	[0.70, 1.74]	.755	.151	Neutral	0.83	[0.53, 1.20]	.815	.143
	Predictive	1.94	[1.15, 2.89]	.997*	.004	Selective	1.11	[0.69, 1.60]	.695	.186

Note. Odds ratio (*OR*) is the odds of target looking for correctly answered items in the offline lexicosemantic knowledge task over the odds of target looking for incorrectly answered items. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; TD = typical language development; DLD = developmental language disorder.

*$p_d > .95$.

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Figure 4. Time course of proportion of target looking in sentences with correct offline lexicosemantic knowledge in Experiment 1. The blue-shaded region denotes the preview period; the red shaded region denotes the predictive period. The vertical solid line denotes start of sentence. Vertical dashed lines indicate the start of linguistically informative material: *The baby is* | *wearing* | *a very special* | *bib*. The horizontal dashed line indicates 50% target and 50% distractor looking. Error bars are standard error of the mean. T / (T + D) = Target proportion / (Target proportion + Distractor proportion). N1 = first noun (agent); N2 = second noun (patient). HL = high–low; HM = high–medium; ML = medium–low.



membership) predict performance in Experiments 1 and 2 to determine whether the differences in language test scores between groups across the two experiments produced exaggerated overall group-level differences, or whether continuous language scores also had similar associations with the offline and online tasks compared to the group-level results. We found highly similar patterns of results in the supplemental analysis compared to the models in the main text, suggesting that differences between experiments were due to stimuli differences and not differences in group composition.

Results

Offline Lexicosemantic Knowledge Task

Figure 6 shows the proportion of trials within each condition that the children selected the higher associated patient in the offline knowledge task. Just as in Experiment 1, the mixed-effects logistic regression model showed strong evidence for differences in accuracy across conditions and between groups (see Table 5). Across groups, there was strong evidence that accuracy in the HL condition, EMM = .97, 95% CI [.94, .99], was higher than in the ML condition, EMM = .91, 95% CI [.84, .97], p_d = .990, p_{ROPE} = .005, but there was weaker evidence that the HL condition and HM condition, EMM = .95, 95% CI [.90, .98], p_d = .909, p_{ROPE} = .040, differed, or that the HM and ML conditions differed, p_d = .893, p_{ROPE} = .045. Across conditions, there was strong evidence that accuracy for the group with TD, EMM = .96, 95% CI [.93, .99], was higher than for the group with DLD, EMM = .93, 95% CI [.88, .97], p_d = .954, p_{ROPE} = .025. In contrast with Experiment 1, there was no evidence for group

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 9. Model results for online predictive processing in sentences with correct offline lexicosemantic knowledge across experiments.

Experiment	Term	Estimate	Error	95% CI	p_d	p_{ROPE}
Experiment 1	Intercept	0.23	0.10	[0.02, 0.43]	.984*	.041
	Condition1	0.09	0.08	[–0.07, 0.24]	.867	.311
	Condition2	–0.01	0.08	[–0.17, 0.15]	.554	.511
	Group	0.03	0.07	[–0.11, 0.16]	.653	.553
	Window	0.26	0.07	[0.11, 0.40]	.999*	.005
	Condition1 × Group	0.00	0.07	[–0.13, 0.13]	.502	.597
	Condition2 × Group	0.05	0.07	[–0.09, 0.18]	.745	.476
	Condition1 × Window	0.06	0.05	[–0.03, 0.15]	.891	.476
	Condition2 × Window	–0.02	0.05	[–0.11, 0.08]	.653	.703
	Group × Window	0.01	0.06	[–0.10, 0.12]	.601	.673
	Condition1 × Group × Window	0.04	0.04	[–0.04, 0.12]	.823	.639
	Condition2 × Group × Window	–0.03	0.04	[–0.12, 0.05]	.768	.675
Experiment 2	Intercept	0.28	0.06	[0.16, 0.40]	1	< .001
	Condition1	0.06	0.07	[–0.07, 0.20]	.822	.405
	Condition2	–0.04	0.07	[–0.17, 0.10]	.691	.522
	Group	0.09	0.05	[–0.01, 0.20]	.961*	.219
	Sentence type	0.06	0.06	[–0.06, 0.18]	.844	.427
	Condition1 × Group	0.14	0.06	[0.02, 0.26]	.989*	.077
	Condition2 × Group	–0.08	0.06	[–0.21, 0.04]	.918**	.302
	Condition1 × Sentence Type	0.02	0.06	[–0.08, 0.13]	.674	.656
	Condition2 × Sentence Type	0.05	0.05	[–0.06, 0.16]	.845	.485
	Group × Sentence Type	0.08	0.06	[–0.04, 0.21]	.917**	.297
	Condition1 × Group × Sentence Type	0.04	0.05	[–0.06, 0.14]	.792	.581
	Condition2 × Group × Sentence Type	–0.01	0.05	[–0.11, 0.09]	.587	.705

Note. Contrast coding was Condition1 = [HL = 1, HM = 0, ML = –1]; Condition2 = [HL = 0, HM = 1, ML = –1]; Group = [TD = 1, DLD = –1]; Window = [Predictive = 1, Preview = –1]; Sentence type = [Selective = 1, Neutral = –1]. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; HL = high–low; HM = high–medium; ML = medium–low; TD = typical language development; DLD = developmental language disorder.

*p_d > .95. **p_d > .90.

differences in the HL condition but stronger evidence for group differences in the HM and ML conditions (see Table 6).

Online Predictive Processing Task

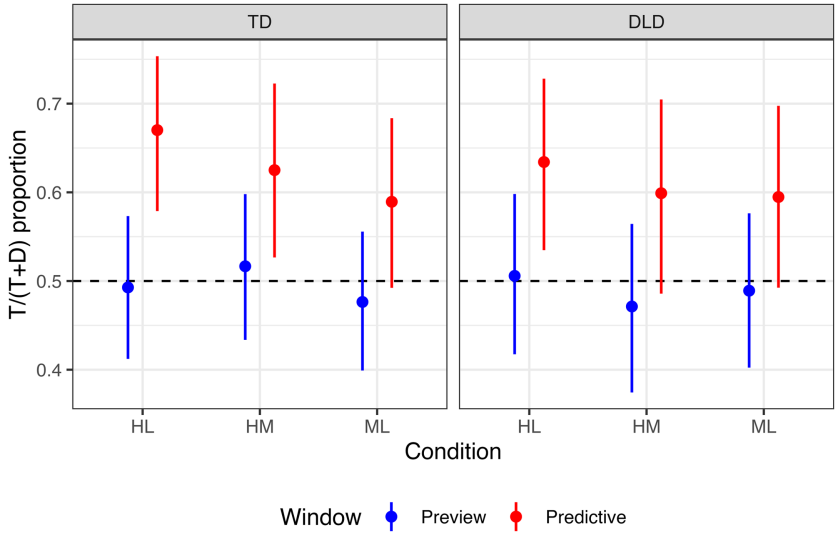
Influence of offline knowledge. Figure 7 shows the proportion of target looking in neutral and selective sentences with correct and incorrect answers on the offline lexicosemantic knowledge task. Just as in Experiment 1, the mixed-effects beta regression model showed strong evidence for the influence of offline accuracy on online predictive processing (see Table 7). There was strong evidence for an interaction between sentence type and offline accuracy; across groups, in selective sentences, items with correct offline accuracy had greater proportions of target looking compared to items with incorrect offline accuracy, OR = 1.32, 95% CI [0.94, 1.74], p_d = .958, p_{ROPE} = .062, while there was little evidence for a difference in neutral sentences, OR = 0.91, 95% CI [0.65, 1.19], p_d = .740, p_{ROPE} = .229. Table 8 and Figure 7 show that the size of the difference between correctly answered versus incorrectly answered items may differ somewhat across groups, with the group with DLD showing little evidence that they predicted more for correctly answered items versus incorrectly answered items in selective sentences. However, we note that Table 7 shows that there was very little evidence for interactions between group and offline accuracy or between group, offline accuracy, and sentence type. The best way to interpret these patterns may be to reserve judgment about group differences and stick to the claim with strong evidence—on average, the children in the sample showed better target looking in selective sentences for correctly answered items.

Online processing with accurate knowledge. Figure 8 shows the time course of target looking in sentences with accurate offline knowledge in the neutral and selective sentences. Table 9 shows the results of the mixed-effects beta regression model examining average target proportion across groups, conditions, and sentence types.

Figure 9 shows the model predictions for target looking across conditions, groups, and sentence types.

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Figure 5. Model-predicted proportion of target looking in sentences with correct offline lexicosemantic knowledge in Experiment 1. Raw participant-level data are presented in Supplemental Figure S3 (see Supplemental Material S1). Error bars are 95% CIs. The horizontal dashed line indicates 50% target and 50% distractor looking. T / (T + D) = Target proportion / (Target proportion + Distractor proportion). TD = typical language development; DLD = developmental language disorder; HL = high–low; HM = high–medium; ML = medium–low.



There was strong evidence for group differences involving sentence type and condition. Specifically, across all conditions and sentence types, children with TD demonstrated higher target looking, EMM = .59, 95% CI [.56, .63], than children with DLD, EMM = .55, 95% CI [.50, .59], p_d = .961, p_{ROPE} = .089. This overall group advantage was modified by interactions with condition. Specifically, Table 10 shows that there was strong evidence for enhanced target looking in the selective versus the neutral sentences for the HL and HM conditions for the group with TD only; the group with DLD showed no evidence for enhanced target looking in any of the conditions. Table 11 shows that there was weak evidence for differences between conditions within groups, with the group with TD having somewhat better target looking in the HL compared to the ML condition. Last, Table 12 shows that there was strong evidence for group differences in the target looking gain in the selective versus the neutral sentences in the HL condition, with the group with TD having enhanced target looking relative to the group with DLD, but there was less evidence for group differences in the other conditions. Overall, the results from this model suggest that the group with DLD showed little predictive looking, and the group with TD showed better predictive looking, particularly in the HL condition.

General Discussion

This study aimed to answer the question: Why do children with DLD have a poorer ability to predict upcoming words in sentences compared to children with TD? By

Table 10. Comparisons for online predictive processing in sentences with correct offline lexicosemantic knowledge across experiments

		Experiment 1				Experiment 2			
Group	Condition	*OR*	95% CI	p_d	p_{ROPE}	*OR*	95% CI	p_d	p_{ROPE}
TD	HL	2.09	[1.28, 3.07]	.999*	< .001	1.53	[0.97, 2.26]	.976*	.028
	HM	1.56	[0.95, 2.32]	.977*	.025	1.46	[0.91, 2.16]	.960*	.045
	ML	1.58	[0.97, 2.38]	.979*	.025	1.08	[0.64, 1.59]	.635	.186
DLD	HL	1.70	[1.05, 2.56]	.990*	.012	0.93	[0.54, 1.40]	.632	.177
	HM	1.68	[0.98, 2.56]	.984*	.017	1.09	[0.62, 1.67]	.639	.171
	ML	1.54	[0.91, 2.33]	.970*	.034	0.86	[0.50, 1.36]	.720	.148

Note. For Experiment 1, odds ratio (*OR*) is the odds of target looking in the predictive window over the odds of target looking in the preview window. For Experiment 2, *OR* is the odds of target looking in selective sentences over the odds of target looking in the neutral sentences. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; TD = typical language development; HL = high–low; HM = high–medium; ML = medium–low; DLD = developmental language disorder.

*p_d > .95.

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 11. Condition differences for online predictive processing in sentences with correct offline lexicosemantic knowledge across experiments.

Group	Comparison	Experiment 1				Experiment 2			
		OR	95% CI	p_d	p_{ROPE}	*OR*	95% CI	p_d	p_{ROPE}
TD	HL - HM	1.34	[0.86, 1.96]	.921*	.079	1.04	[0.62, 1.58]	.573	.179
	HL - ML	1.33	[0.83, 1.95]	.907*	.089	1.42	[0.81, 2.16]	.924*	.066
	HM - ML	0.99	[0.61, 1.46]	.522	.198	1.35	[0.77, 2.09]	.891	.085
DLD	HL - HM	1.01	[0.59, 1.52]	.517	.191	0.85	[0.46, 1.36]	.733	.137
	HL - ML	1.10	[0.66, 1.65]	.670	.178	1.07	[0.55, 1.73]	.593	.152
	HM - ML	1.09	[0.64, 1.68]	.645	.168	1.26	[0.65, 2.07]	.790	.112

Note. For Experiment 1, odds ratio (*OR*) is the difference between conditions in the odds of target looking in the predictive window over the odds of target looking in the preview window. For Experiment 2, *OR* is the difference between conditions in the odds of target looking in selective sentences over the odds of target looking in the neutral sentences. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; TD = typical language development; HL = high–low; HM = high–medium; ML = medium–low; DLD = developmental language disorder.

*$p_d > .90$.

comparing results across offline knowledge and online processing tasks in sentences with different speaking rates, speeds, and complexities, we found an answer: Both lexicosemantic knowledge and processing factors help explain why children with DLD have poorer sentence prediction ability (see Table 1 for the summary of results). The children with DLD had poor offline lexicosemantic knowledge, which resulted in less accurate online sentence prediction. In addition, even after controlling for lexicosemantic knowledge, the children with DLD still showed less accurate sentence prediction, implicating processing factors involving working memory, sustained attention, and online activation/inhibition of verb semantic features.

Lexicosemantic Knowledge

Across the two experiments, the children with DLD had poorer lexicosemantic knowledge compared to their peers with TD. In an offline task, the children with DLD pointed to the most associated patient for a verb/agent combination (e.g., bib vs. necklace for wear/baby) less often than the children with TD. This pattern fits with prior work, showing that these children have less breadth and depth in lexicosemantic knowledge (e.g., McGregor et al., 2002, 2013) but extends these findings to the semantics of verbs' arguments. Across both experiments, compared to the group with TD, the group with DLD had more difficulty making distinctions between alternatives in all conditions. However, this difficulty was most consistently observed in the HM condition, a condition requiring more fine- versus coarse-grained distinctions between alternatives (e.g., *bib* vs. *sweater* compared to *bib* vs. *necklace for wear/baby*). This pattern suggests that children with DLD may have lexicosemantic knowledge deficits that prevent accurate computation of small differences in associative strength. These deficits may include less robust representations of semantic features or difficulties using the full suite of semantic features in comparisons (compared to using just the most relevant semantic features, which might match the verb/agent combination broadly but fail to discriminate among options close in associative strength). This explanation is consistent with work like McMurray et al. (2014) showing that individuals with poorer language skills are slower and less accurate in lexical identification tasks when targets and distractors share semantic features.

Table 12. Group differences for online predictive processing in sentences with correct offline lexicosemantic knowledge across experiments.

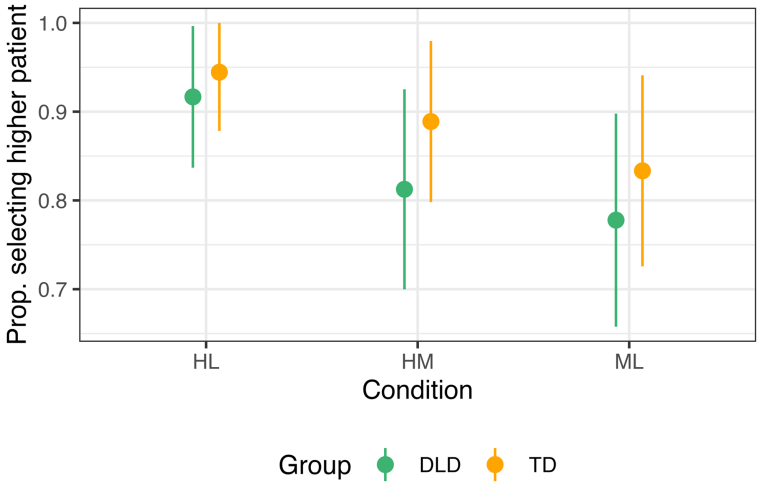
Comparison	Condition	Experiment 1				Experiment 2			
		OR	95% CI	p_d	p_{ROPE}	*OR*	95% CI	p_d	p_{ROPE}
TD vs. DLD	HL	1.24	[0.66, 2.00]	.781	.118	1.65	[0.78, 2.84]	.950*	.038
	HM	0.93	[0.49, 1.55]	.604	.149	1.34	[0.60, 2.33]	.821	.088
	ML	1.03	[0.52, 1.68]	.542	.154	1.25	[0.54, 2.19]	.749	.106

Note. For Experiment 1, odds ratio (*OR*) is the difference in typical language development (TD) versus developmental language disorder (DLD) in the odds of target looking in the predictive window over the odds of target looking in the preview window. For Experiment 2, *OR* is the difference in TD versus DLD in the odds of target looking in selective sentences over the odds of target looking in the neutral sentences. CI = confidence interval; p_d = probability of direction; p_{ROPE} = proportion of posterior samples in region of practical equivalence; HL = high–low; HM = high–medium; ML = medium–low.

*$p_d > .95$.

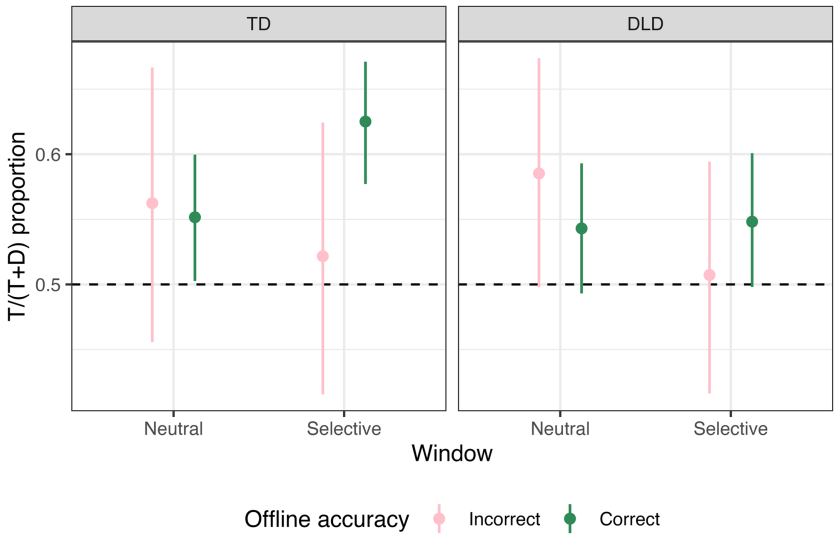
Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Figure 6. Accuracy on offline lexicosemantic knowledge task in Experiment 2. Raw participant-level data are presented in Supplemental Figure S4 (see Supplemental Material S1). Error bars are 95% confidence intervals. Prop. = proportion; HL = high–low; HM = high–medium; ML = medium–low; DLD = developmental language disorder; TD = typical language development.



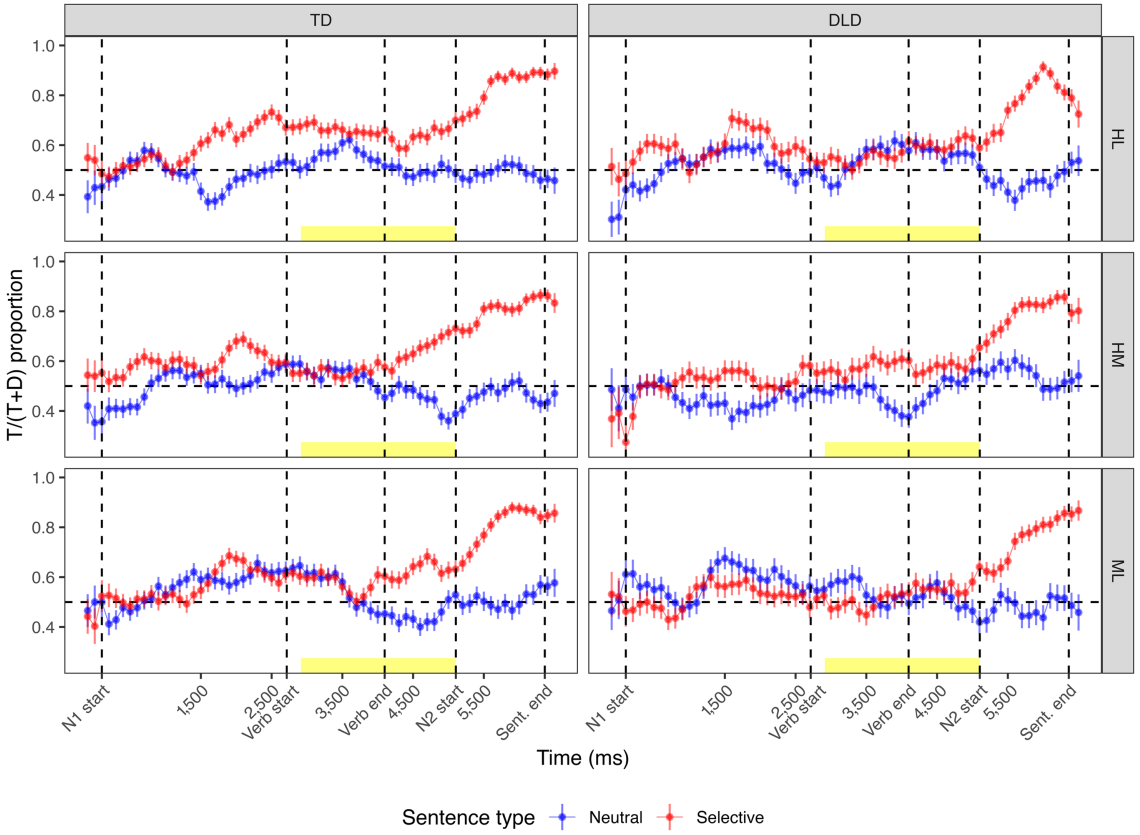
The results provided strong evidence for Hypothesis 1, namely, that lexicosemantic knowledge deficits at the verb argument level impact the robustness of sentence prediction in children with DLD. When children demonstrated knowledge of patient–verb–agent associates in the offline lexicosemantic knowledge task, they demonstrated better sentence prediction of those specific patients in the online eye tracking task. Given that the children with DLD had poorer offline lexicosemantic knowledge, they also had poorer overall online sentence prediction compared to their peers. While prior work suggested that lexicosemantic knowledge deficits could explain differences between children with DLD and peers with TD (Andreu et al., 2013; Borovsky et al., 2013), ours is the first study to explicitly test the hypothesis that the deficits seen in sentence prediction in sentences using early-developing verbs arise, at least in part, from poor lexicosemantic knowledge. This finding offers a clear clinical implication: When teaching a verb, the semantics of the verb's arguments—not just the semantics of the verb itself—should be targeted. Work on toddlers with TD and late talkers suggests that in some situations, teaching verbs with a

Figure 7. Model-predicted proportion of target looking for sentences with incorrect versus correct offline knowledge in Experiment 2. Raw participant-level data are presented in Supplemental Figure S5 (see Supplemental Material S1). The horizontal dashed line indicates 50% target and 50% distractor looking. Error bars are 95% confidence intervals. T / (T + D) = Target proportion / (Target proportion + Distractor proportion). TD = typical language development; DLD = developmental language disorder.



Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Figure 8. Time course of proportion of target looking in sentences with correct offline lexicosemantic knowledge in Experiment 2. Blue denotes neutral (nonsemantically selective) sentences; red denotes semantically selective sentences. Yellow region shows predictive window. Vertical dashed lines indicate the start of linguistically informative material: *The baby is* | *wearing* | *a* | *bib*. The horizontal dashed line indicates 50% target and 50% distractor looking. Error bars are standard error of the mean. T / (T + D) = Target proportion / (Target proportion + Distractor proportion). N1 = first noun (agent); N2 = second noun (patient); TD = typical language development; DLD = developmental language disorder; HL = high–low; HM = high–medium; ML = medium–low.



variety of associated agents/patients (e.g., Childers et al., 2016; Haryu et al., 2011; Kueser, Peters, & Borovsky, 2023) and in different kinds of syntactic structures that vary argument roles (Nicholas et al., 2024) can help verb learning; these procedures help verb learning likely because they highlight the semantic content of verbs' argument positions. Thus, focusing on the semantics of verb argument positions may not only improve overall verb lexicosemantic knowledge, which may then support the learning of other verbs (Kueser, Horvath, & Borovsky, 2023; Saji et al., 2011), but may also improve the speed and accuracy of sentence processing.

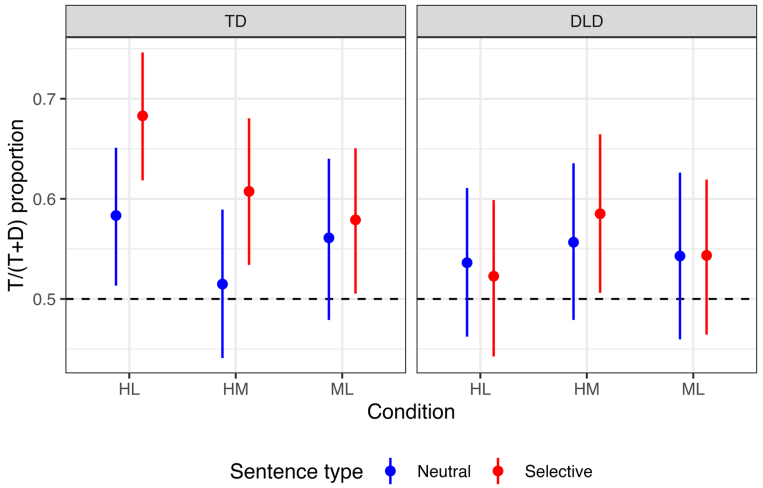
Processing Factors

We also asked whether processing factors explain patterns of sentence prediction by children with DLD. We found that even after accounting for item-level lexicosemantic knowledge, the children with DLD differed from their peers with TD. Specifically, sentence prediction in a fast sentence was possible for the children with DLD, but in slower sentences, their predictions appeared to decay over time.

This pattern of performance offers strong evidence for Hypothesis 2B, namely, that the poorer sentence prediction experienced by children with DLD is explained by problems with quick decay of material in working memory and/or difficulties with sustained attention. In contrast, given that the children with DLD showed better performance in fast sentences, there was little evidence for Hypothesis 2A's proposal for a contribution of slower processing speed or increased interference in working memory. In fact, we note that even in the faster sentences in Experiment 1, where average target looking was similar in the two groups, the group with DLD demonstrated a qualitatively rougher looking trajectory compared to the group with TD, again suggesting that they had difficulty maintaining attention on their predictions and/or keeping predictions active in working memory. In addition, this pattern qualitatively mirrors some of the time course data

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Figure 9. Model-predicted proportion of target looking in sentences with correct offline lexicosemantic knowledge in Experiment 2. Raw participant-level data are presented in Supplemental Figure S6 (see Supplemental Material S1). Error bars are 95% confidence intervals. The horizontal dashed line indicates 50% target and 50% distractor looking. T / (T + D) = Target proportion / (Target proportion + Distractor proportion). TD = typical language development; DLD = developmental language disorder; HL = high–low; HM = high–medium; ML = medium–low.



from prior sentence prediction studies in DLD, which reveal a much less even upward trajectory of target looking for individuals with DLD compared to individuals with TD (e.g., Borovsky et al., 2013; Kueser et al., 2024), with numerous discontinuities in predictions across time, consistent with memory decay or difficulties with sustained attention.

More evidence for the idea that memory decay and/or sustained attention difficulties influenced the group with DLD comes from a supplemental analysis of Experiment 2 data (see Supplemental Material S1). In this analysis, we examined target looking in a window from the agent to the onset of the verb (e.g., "The baby is wearing . . . "), motivated by the visual impression that predictive looking for the group with DLD was about as robust as their peers' early in the sentence and grew less robust as the sentence unfolded. Just as in the main analysis focused on prediction after the verb, we found that for the group with TD, there was evidence for prediction in the HL and HM conditions. Surprisingly, we also found evidence for prediction for the group with DLD in the HL condition, despite there being no evidence for prediction for this group in the main analysis. Again, this pattern suggests working memory decay and sustained attention deficits—children with DLD may have made a prediction earlier in the sentence but had difficulty holding onto it. This finding is very similar to findings from a study on sentence prediction in passives (Kueser et al., 2024). In that study, predictions of sentence-final agents in passive sentences (e.g., "The box was pushed very hard by the woman") were measured in 4- to 5-year-old children with and without DLD. The group with DLD showed similar patterns of sentence prediction from the onset of the verb to the onset of the *by*-phrase; however, upon hearing "by," the group with DLD quickly diverged from the group with TD, showing much increased distractor looking while the group with TD continued to show increases in correct prediction. That pattern of performance was interpreted as evidence that children with DLD have difficulty integrating redundant information (i.e., verb semantics with sentence syntax) to make a stronger prediction. This interpretation also fits the data here—namely, that despite making similarly correct predictions early on in the sentence, the group with DLD failed to use verb semantics to solidify their predictions later on in the sentence because their memory for their earlier prediction decayed and/or their attention wandered.

Performance across conditions revealed additional patterns suggesting that graded prediction may also be a difficulty for children with DLD, consistent with Hypothesis 2C. The strongest evidence for graded prediction deficits came from group differences that varied across conditions in Experiment 2. In that experiment, there was strong evidence that the two groups differed in the HL condition but little evidence for differences in the HM or ML conditions; this is despite the fact that there was little evidence that the groups differed in offline lexicosemantic knowledge for the HL condition. On one hand, this pattern can be interpreted as evidence for the perspective that, despite having strong enough lexicosemantic knowledge of the patients that should go with the verbs, the children with DLD could not deploy that strong knowledge in

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

online prediction. Thus, children with DLD may have difficulty appropriately weighting semantic information to engage in online prediction. On the other hand, given that the children with DLD showed no evidence of sentence prediction at all in any of the conditions in Experiment 2, group differences across conditions may be better interpreted as evidence that the group with TD had the best prediction in the HL condition and that the task may have been uniformly too difficult for the group with DLD. In fact, there was no evidence for group differences across conditions in Experiment 1, suggesting that, in contexts where children with DLD can make robust sentence predictions, their ability to do so in a graded manner is largely similar to their peers'. Thus, these children's graded prediction deficits may have a relatively smaller impact on sentence prediction compared to working memory decay and sustained attention deficits and may arise only in difficult processing contexts. Furthermore, given that the offline lexicosemantic knowledge tasks showed an opposite pattern in which the most consistent group differences were in the HM condition (where differences in associative strength were smaller than in the HL condition), it may be too soon to describe with certainty how graded lexicosemantic processing works in children with DLD.

Clinically, these findings suggest the use of supports for working memory and attention, including visuals that could help reactivate memory traces or strategies to keep children's attention engaged. Further work will be needed to determine how such supports might translate to better sentence comprehension in real-life contexts. The results also suggest that language directed to children with DLD should be presented at a normal rate rather than a slowed rate to reduce the effect of working memory decay and difficulties with sustained attention. However, this suggestion is inconsistent with two prior studies examining how speaking rate impacts language comprehension in DLD (Montgomery, 2004, 2005). Potential reasons for this inconsistency across studies include differences in measures (e.g., eye tracking measures relating to sentence prediction in this study vs. picture-pointing measures relating to overall sentence comprehension), differences in the types of sentences used (e.g., relatively simple active sentences in this study vs. more complex sentences), and differences how rate was manipulated across studies (e.g., simultaneous changes in speaking rate, sentence length, and sentence complexity in this study vs. changes in only speaking rate). More work is needed to determine how these factors together influence the effect of speaking rate on children with DLD.

Limitations and Future Directions

This study represents a first step in understanding how sentence prediction in children with DLD differs from peers with TD. There are at least three outstanding issues for future work to address, however.

First, it is not clear what aspects of lexicosemantic knowledge are difficult for children with DLD. This study's stimuli conflated two pieces of information likely to be important—multiword cooccurrence statistics (Abu-Zhaya et al., 2022; Onnis & Huettig, 2021) and verb-specific semantic features (Altmann & Kamide, 1999; Fernald et al., 2008; Gambi et al., 2021). Multiword cooccurrence statistics reflect the frequency of "chunks" in the linguistic input and are likely important for processing sentences often heard in the environment. Verb-specific semantic features detail information about event participants such as agents and patients (e.g., a patient of the verb "dry" may be "clothing" or something <wet>; McRae et al., 1997; Metusalem et al., 2012) and are likely important for processing novel sentences. Given that there are differences in the semantic properties of verbs known by toddlers with language delays compared to peers with TD (Horvath et al., 2019, 2022), it is possible that verb-specific semantic feature knowledge may be particularly affected relative to multiword co-occurrence statistics knowledge, causing difficulty processing novel sentences.

Second, while comparisons between experiments provided valuable information allowing inferences about the effect of cognitive differences between groups on sentence prediction, directly measuring the cognitive mechanisms explored here—working memory capacity/robustness, processing speed, sustained attention, and so forth—and relating individual differences in these measures to individual differences in sentence prediction ability would create a stronger case. Additionally, understanding of these relationships would be helped by a systematic exploration of how sentence-level characteristics affect sentence prediction. Specifically, this study's stimuli conflated sentence rate, sentence length, and syntactic complexity; teasing apart these characteristics would be informative to better understand how different cognitive mechanisms in DLD are implicated.

Third, our conclusions were based on between-participants comparisons across experiments. Despite using the same inclusionary and exclusionary criteria for enrolment in both experiments, with respect to language test scores, the group with TD differed from the group with DLD slightly more in Experiment 1 compared to Experiment 2. Though an analysis presented in Supplemental Material S1 using continuous language test scores rather than group membership mirrored the results presented in the main text, differences in samples between experiments can complicate comparisons between them. Future work could address this issue by examining these patterns within a single sample.

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Conclusions

Prediction has been argued to be omnipresent in language and in cognition more generally (Altmann & Mirković, 2009; DeLong et al., 2014; Kukona et al., 2016; Kuperberg & Jaeger, 2016; Pickering & Garrod, 2013). Theories suggest that prediction speeds language processing in the moment, induces a state of preparedness to interpret language, helps fine-tune the language system over time, and provides valuable error signals for learning (Bannon et al., 2025; Dell & Chang, 2014; Ferreira & Chantavarin, 2018; Levy, 2008; Pickering & Garrod, 2013). Prior work on prediction abilities in children with DLD had often showed differences from same-age peers with TD, with children with DLD being slower and less accurate in predicting the final word in sentences. However, the two major potential reasons for these differences—poorer lexicosemantic abilities and poorer online processing—had not been separately explored. The current study provided a first attempt to distinguish between these mechanisms, finding that children with DLD have sentence prediction deficits due to a combination of lexicosemantic knowledge deficits and problems with working memory decay and sustained attention.

Data Availability Statement

Data and analysis scripts are available at https://osf.io/cmesb/.

Acknowledgments

Justin B. Kueser, Arielle Borovsky, and Claney Outzen were supported by National Institute on Deafness and Other Communication Disorders (NIDCD) Grant R01 DC018593 to Arielle Borovsky. Pat Deevy and Laurence B. Leonard were supported by NIDCD Grant R01 DC014708 to Laurence B. Leonard. Justin B. Kueser was additionally supported by NIDCD Grant R21 DC022371.

References

Abu-Zhaya, R., Arnon, I., & Borovsky, A. (2022). Do children use multi-word information in real-time sentence comprehension? *Cognitive Science, 46*(3), Article e13111. https://doi.org/10.1111/cogs.13111

Alt, M., Plante, E., & Creusere, M. (2004). Semantic features in fast-mapping: Performance of preschoolers with specific language impairment versus preschoolers with normal language. *Journal of Speech, Language, and Hearing Research, 47*(2), 407–420. https://doi.org/10.1044/1092-4388(2004/033)

Alt, M., & Spaulding, T. (2011). The effect of time on word learning: An examination of decay of the memory trace and vocal rehearsal in children with and without specific language impairment. *Journal of Communication Disorders, 44*(6), 640–654. https://doi.org/10.1016/j.jcomdis.2011.07.001

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33*(4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x

Andreu, L., Sanz-Torrent, M., & Trueswell, J. C. (2013). Anticipatory sentence processing in children with specific language impairment: Evidence from eye movements during listening. *Applied Psycholinguistics, 34*(1), 5–44. https://doi.org/10.1017/S0142716411000592

Bannon, J., Gollan, T. H., & Ferreira, V. S. (2025). Is predicting during language processing worth it? Effects of cloze probability and semantic similarity on failed predictions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 51*(1), 106–118. https://doi.org/10.1037/xlm0001347

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *The Journal of Child Psychology and Psychiatry, 37*(4), 391–403. https://doi.org/10.1111/j.1469-7610.1996.tb01420.x

Borovsky, A. (2022). Developmental changes in how children generalize from their experience to support predictive linguistic processing. *Journal of Experimental Child Psychology, 219,* Article 105349. https://doi.org/10.1016/j.jecp.2021.105349

Borovsky, A., Burns, E., Elman, J. L., & Evans, J. L. (2013). Lexical activation during sentence comprehension in adolescents with history of specific language impairment. *Journal of Communication Disorders, 46*(5–6), 413–427. https://doi.org/10.1016/j.jcomdis.2013.09.001

Borovsky, A., Elman, J. L., & 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*(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), Article 395. https://doi.org/10.32614/RJ-2018-017

Carter, B. T., Foster, B., Muncy, N. M., & Luke, S. G. (2019). Linguistic networks associated with lexical, semantic and syntactic predictability in reading: A fixation-related fMRI study. *NeuroImage, 189,* 224–240. https://doi.org/10.1016/j.neuroimage.2019.01.018

Childers, J. B., Parrish, R., Olson, C. V., Burch, C., Fung, G., & McIntyre, K. P. (2016). Early verb learning: How do children learn how to compare events? *Journal of Cognition and Development, 17*(1), 41–66. https://doi.org/10.1080/15248372.2015.1042580

Childers, J. B., & Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 311–335). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195170009.003.0013

Cohen, J. (1988). Set correlation and contingency tables. *Applied Psychological Measurement, 12*(4), 425–434. https://doi.org/10.1177/014662168801200410

Dawson, J., Stout, C., Eyer, J., Tattersall, P., Fonkalsrud, J., & Croley, K. (2005). *Structured Photographic Expressive Language Test–Preschool 2*. Janelle Publications. https://www.janellepublications.com/0707.shtml

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition.

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1634), Article 20120394. https://doi.org/10.1098/rstb.2012.0394

DeLong, K. A., Troyer, M., & Kutas, M. (2014). Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. *Language and Linguistics Compass, 8*(12), 631–645. https://doi.org/10.1111/lnc3.12093

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136–1146. https://doi.org/10.1044/jslhr.4105.1136

Dunn, D. M. (2019). *Peabody Picture Vocabulary Test–Fifth Edition* [Computer software]. NCS Pearson. https://www.pearsonassessments.com/en-us/Store/Professional-Assessments/Academic-Learning/Peabody-Picture-Vocabulary-Test-%7C-Fifth-Edition/p/100001984

Ebert, K. D., & Kohnert, K. (2011). Sustained attention in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 54*(5), 1372–1384. https://doi.org/10.1044/1092-4388(2011/10-0231)

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholonguistics: On-line methods in children's language processing* (pp. 97–135). John Benjamins. https://doi.org/10.1075/lald.44.06fer

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science, 27*(6), 443–448. https://doi.org/10.1177/0963721418794491

Ferreira, F., & Ferreira, V. S. (2024). Psycholinguistics. In *Open encyclopedia of cognitive science* (1st ed.). MIT Press. https://doi.org/10.21428/e2759450.8207aa16

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 52*(4), 915–929. https://doi.org/10.1044/1092-4388(2009/07-0053)

Gambi, C., Jindal, P., Sharpe, S., Pickering, M. J., & Rabagliati, H. (2021). The relation between preschoolers' vocabulary development and their ability to predict and recognize words. *Child Development, 92*(3), 1048–1066. https://doi.org/10.1111/cdev.13465

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Erlbaum.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.010

Gillam, R. B., Cowan, N., & Day, L. S. (1995). Sequential memory in children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 38*(2), 393–402. https://doi.org/10.1044/jshr.3802.393

Gillam, R. B., Cowan, N., & Marler, J. A. (1998). Information processing by school-age children with specific language impairment: Evidence from a modality effect paradigm. *Journal of Speech, Language, and Hearing Research, 41*(4), 913–926. https://doi.org/10.1044/jslhr.4104.913

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1*(1), 23–64. https://doi.org/10.1207/s15473341lld0101_4

Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test–Preschool: Second Edition. *Language, Speech, and Hearing Services in Schools, 40*(2), 150–160. https://doi.org/10.1044/0161-1461(2008/07-0049)

Hadley, P. A. (2020). Exploring sentence diversity at the boundary of typical and impaired language abilities. *Journal of Speech, Language, and Hearing Research, 63*(10), 3236–3251. https://doi.org/10.1044/2020_JSLHR-20-00031

Hadley, P. A., McKenna, M. M., & Rispoli, M. (2018). Sentence diversity in early language development: Recommendations for target selection and progress monitoring. *American Journal of Speech-Language Pathology, 27*(2), 553–565. https://doi.org/10.1044/2017_AJSLP-17-0098

Hadley, P. A., Rispoli, M., & Hsu, N. (2016). Toddlers' verb lexicon diversity and grammatical outcomes. *Language, Speech, and Hearing Services in Schools, 47*(1), 44–58. https://doi.org/10.1044/2015_LSHSS-15-0018

Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extension. *Child Development, 82*(2), 674–686. https://doi.org/10.1111/j.1467-8624.2010.01567.x

Horvath, S., Kueser, J. B., Kelly, J., & Borovsky, A. (2022). Difference or delay? Syntax, semantics, and verb vocabulary development in typically developing and late-talking toddlers. *Language Learning and Development, 18*(3), 352–376. https://doi.org/10.1080/15475441.2021.1977645

Horvath, S., Rescorla, L., & Arunachalam, S. (2019). The syntactic and semantic features of two-year-olds' verb vocabularies: A comparison of typically developing children and late talkers. *Journal of Child Language, 46*(3), 409–432. https://doi.org/10.1017/S0305000918000508

Huettig, F., Audring, J., & Jackendoff, R. (2022). A parallel architecture perspective on pre-activation and prediction in language processing. *Cognition, 224,* Article 105050. https://doi.org/10.1016/j.cognition.2022.105050

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77*(6), 1822–1841. https://doi.org/10.1111/j.1467-8624.2006.00976.x

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 53*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kaufman, A. S., & Kaufman, N. L. (2018). *Kaufman Assessment Battery for Children–Second Edition normative update* [Computer software]. Pearson. https://www.pearsonassessments.com/en-us/Store/Professional-Assessments/Cognition-%26-Neuro/Kaufman-Assessment-Battery-for-Children-%7C-Second-Edition-Normative-Update/p/100000088

Kay, M. (2021). *tidybayes: Tidy data and geoms for Bayesian models*. https://doi.org/10.5281/zenodo.1308151

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review, 25*(1), 178–206. https://doi.org/10.3758/s13423-016-1221-4

Kueser, J. B., & Borovsky, A. (2025). Complex meanings shape early noun and verb vocabulary structure and learning. Canadian *Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale, 79*(1), 85–97. https://doi.org/10.1037/cep0000355

Kueser, J. B., Borovsky, A., Deevy, P., Muezzinoglu, M., Outzen, C., & Leonard, L. B. (2024). Verb vocabulary supports event probability use in developmental language disorder. *Journal of Speech, Language, and Hearing Research, 67*(5), 1490–1513. https://doi.org/10.1044/2024_JSLHR-23-00600

Kueser, J. B., Horvath, S., & Borovsky, A. (2023). Two pathways in vocabulary development: Large-scale differences in noun

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

and verb semantic structure. *Cognitive Psychology, 143,* Article 101574. https://doi.org/10.1016/j.cogpsych.2023.101574

Kueser, J. B., Peters, R., & Borovsky, A. (2023). The role of semantic similarity in verb learning events: Vocabulary-related changes across early development. *Journal of Experimental Child Psychology, 226,* Article 105565. https://doi.org/10.1016/j.jecp.2022.105565

Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., Pugh, K. R., Shankweiler, D. P., & Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica, 171,* 72–84. https://doi.org/10.1016/j.actpsy.2016.09.009

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software, 69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50*(2), 408–428. https://doi.org/10.1044/1092-4388(2007/029)

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology, 88,* 22–60. https://doi.org/10.1016/j.cogpsych.2016.06.002

MacWhinney, B. (2015). Psycholinguistics: Overview. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 353–358). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.57010-X

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software, 4*(40), Article 1541. https://doi.org/10.21105/joss.01541

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843–847. https://doi.org/10.1037/a0029284

Marton, K., Campanelli, L., Eichorn, N., Scheuer, J., & Yoon, J. (2014). Information processing and proactive interference in children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*(1), 106–119. https://doi.org/10.1044/1092-4388(2013/12-0306)

Marton, K., Eichorn, N., Campanelli, L., & Zakarias, L. (2016). Working memory and interference control in children with specific language impairment. *Language and Linguistics Compass, 10*(5), 211–224. https://doi.org/10.1111/lnc3.12189

McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of Autism and Developmental Disorders, 42*(1), 35–47. https://doi.org/10.1007/s10803-011-1210-4

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45*(5), 998–1014. https://doi.org/10.1044/1092-4388(2002/081)

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language & Communication Disorders, 48*(3), 307–319. https://doi.org/10.1111/1460-6984.12008

McMurray, B., Munson, C., & Tomblin, J. B. (2014). Individual differences in language ability are related to variation in word recognition, not speech perception: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research, 57*(4), 1344–1362. https://doi.org/10.1044/2014_JSLHR-L-13-0196

McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology, 60*(1), 1–39. https://doi.org/10.1016/j.cogpsych.2009.06.003

McRae, K., Ferretti, T., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes, 12*(2–3), 137–176. https://doi.org/10.1080/016909697386835

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language, 66*(4), 545–567. https://doi.org/10.1016/j.jml.2012.01.001

Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(2), 416–433. https://doi.org/10.1044/1092-4388(2001/034)

Miller, C. A., Leonard, L. B., Kail, R. V., Zhang, X., Tomblin, J. B., & Francis, D. J. (2006). Response time in 14-year-olds with language impairment. *Journal of Speech, Language, and Hearing Research, 49*(4), 712–728. https://doi.org/10.1044/1092-4388(2006/052)

Montgomery, J. W. (2004). Sentence comprehension in children with specific language impairment: Effects of input rate and phonological working memory. *International Journal of Language & Communication Disorders, 39*(1), 115–133. https://doi.org/10.1080/13682820310001616985

Montgomery, J. W. (2005). Effects of input rate and age on the real-time language processing of children with specific language impairment. *International Journal of Language & Communication Disorders, 40*(2), 171–188. https://doi.org/10.1080/13682820400011069

Montgomery, J. W., & Windsor, J. (2007). Examining the language performances of children with and without specific language impairment: Contributions of phonological short-term memory and speed of processing. *Journal of Speech, Language, and Hearing Research, 50*(3), 778–797. https://doi.org/10.1044/1092-4388(2007/054)

Naigles, L.(1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357–374. https://doi.org/10.1017/S0305000900013817

Nicholas, K., Grierson, T., Helen, P., Miller, C., & Van Horne, A. O. (2024). Varying syntax to enhance verb-focused intervention for 30-month-olds with language delay: A concurrent multiple baseline design. *Journal of Speech, Language, and Hearing Research, 67*(2), 562–572. https://doi.org/10.1044/2023_JSLHR-23-00398

Oberauer, K., & Lewandowsky, S. (2014). Further evidence against decay in working memory. *Journal of Memory and Language, 73,* 15–30. https://doi.org/10.1016/j.jml.2014.02.003

Onnis, L., & Huettig, F. (2021). Can prediction and retrodiction explain whether frequent multi-word phrases are accessed 'precompiled' from memory or compositionally constructed on the fly? *Brain Research, 1772,* Article 147674. https://doi.org/10.1016/j.brainres.2021.147674

Park, J. S., Miller, C. A., Sanjeevan, T., van Hell, J. G., Weiss, D. J., & Mainela-Arnold, E. (2020). Bilingualism and processing speed in typically developing children and children with developmental language disorder. *Journal of Speech, Language, and Hearing Research, 63*(5), 1479–1493. https://doi.org/10.1044/2020_JSLHR-19-00403

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Poll, G. H., Watkins, H. S., & Miller, C. A. (2014). Lexical decay during online sentence processing in adults with specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*(6), 2253–2260. https://doi.org/10.1044/2014_JSLHR-L-13-0265

R Development Core Team. (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org

Reuter, T., Mazzei, C., Lew-Williams, C., & Emberson, L. (2023). Infants' lexical comprehension and lexical anticipation abilities are closely linked in early language development. *Infancy, 28*(3), 532–549. https://doi.org/10.1111/infa.12534

Saji, N., Imai, M., Saalbach, H., Zhang, Y., Shu, H., & Okada, H. (2011). Word learning does not end at fast-mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition, 118*(1), 45–61. https://doi.org/10.1016/j.cognition.2010.09.007

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8*(4), 448–467. https://doi.org/10.1037/1082-989X.8.4.448

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale–Second Edition* [Computer software]. Western Psychological Services. https://www.wpspublish.com/cars-2-childhood-autism-rating-scale-second-edition.html

Seiger-Gardner, L., & Brooks, P. J. (2008). Effects of onset- and rhyme-related distractors on phonological processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 51*(5), 1263–1281. https://doi.org/10.1044/1092-4388(2008/07-0079)

Seiger-Gardner, L., & Schwartz, R. G. (2008). Lexical access in children with and without specific language impairment: A cross-modal picture-word interference study. *International Journal of Language & Communication Disorders, 43*(5), 528–551. https://doi.org/10.1080/13682820701768581

Smolak, E., McGregor, K. K., Arbisi-Kelm, T., & Eden, N. (2020). Sustained attention in developmental language disorder and its relation to working memory and language. *Journal of Speech, Language, and Hearing Research, 63*(12), 4096–4108. https://doi.org/10.1044/2020_JSLHR-20-00265

Spaulding, T. J. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(3), 725–738. https://doi.org/10.1044/1092-4388(2009/09-0041)

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass, 9*(8), 311–327. https://doi.org/10.1111/lnc3.12151

Szewczyk, J. M., & Federmeier, K. D. (2022). Context-based facilitation of semantic access follows both logarithmic and linear functions of stimulus probability. *Journal of Memory and Language, 123,* Article 104311. https://doi.org/10.1016/j.jml.2021.104311

van Alphen, P., Brouwer, S., Davids, N., Dijkstra, E., & Fikkert, P. (2021). Word recognition and word prediction in preschoolers with (a suspicion of) a developmental language disorder: Evidence from eye tracking. *Journal of Speech, Language, and Hearing Research, 64*(6), 2005–2021. https://doi.org/10.1044/2021_JSLHR-20-00227

Williams, K. (2019). *Expressive Vocabulary Test–Third Edition* [Computer software]. NCS Pearson. https://www.pearsonassessments.com/en-us/Store/Professional-Assessments/Academic-Learning/Expressive-Vocabulary-Test-%7C-Third-Edition/p/100001982

Witherstone, H. (2024). Speed of processing in developmental language disorder (DLD): The case of real-time grammatical processing. *International Journal of Language & Communication Disorders, 59*(4), 1489–1504. https://doi.org/10.1111/1460-6984.13014

Yazbec, A., Kaschak, M. P., & Borovsky, A. (2019). Developmental timescale of rapid adaptation to conflicting cues in real-time sentence processing. *Cognitive Science, 43*(1), Article e12704. https://doi.org/10.1111/cogs.12704

Zapparrata, N. M., Brooks, P. J., & Ober, T. (2023). Developmental language disorder is associated with slower processing across domains: A meta-analysis of time-based tasks. *Journal of Speech, Language, and Hearing Research, 66*(1), 325–346. https://doi.org/10.1044/2022_JSLHR-22-00221

Downloaded from: https://pubs.asha.org 192.88.142.200 on 11/16/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions