Research Article

Deficits in the Use of Verb Bias Information in Real-Time Processing by College Students With Developmental Language Disorder

Jessica E. Hall,^a Amanda Owen Van Horne,^b
Karla K. McGregor,^{a,c} and Thomas A. Farmer^{a,d}

Purpose: This study examined whether college students with developmental language disorder (DLD) showed similar sensitivity to verb bias information during real-time sentence processing as typically developing (TD) peers.

Method: Seventeen college students with DLD and 16 TD college students participated in a mouse-tracking experiment that utilized the visual world paradigm to examine real-time sentence processing. In experimental trials, participants chose 1 of 2 pictured interpretations of a sentence. Measures of interest were the choice of interpretation and the amount

of competition from the unchosen picture as measured by mouse curvature.

Results: Choice of interpretation and mouse movements by college students with DLD suggested less sensitivity to verb bias information than their TD peers.

Conclusion: College students with DLD showed less evidence of sensitivity to verb bias information than their TD peers in this task. Their performance may reflect the use of compensatory processing strategies and may be related to poor comprehension abilities often observed in this population.

To comprehend spoken language in real time, the brain does some predicting along the way. Imagine hearing the beginning of a sentence, “He said he wants—.” Probability-based expectations can arise from the pragmatic or semantic context of the conversation and from the words within the sentence, particularly from the verb. The verb *want* is quite versatile: It can be used in different syntactic constructions, some more likely (e.g., *wants* + nonfinite clauses such as *to take the subway* or *wants* + direct objects such as *salad*) and others less so (e.g., *wants* + prepositional phrases such as *for nothing*). Less versatile verbs are more predictable: For example, in the sentence, “He said he slept—”, the verb *sleep* can appear in fewer syntactic constructions and is more limited in meaning; thus, the number of possible following words and interpretations is

substantially narrower. Using this type of distributional information during language processing allows for speedy, efficient, and effortless comprehension for competent adult language users. For some individuals, such as those with developmental language disorder (DLD), language processing is disrupted and comprehension is impaired. Focusing on the real-time process of comprehension in DLD could help identify how breakdowns in comprehension occur. Individuals with DLD hear similar amounts of language as their peers without disorders (see Leonard, 2014), but their actual experience may differ greatly if they are not processing the information they hear in the same way. By employing online processing techniques such as eye or mouse tracking with this population, we may identify factors that contribute to poorer overall comprehension.

Language Processing

Studies of sentence processing are based on the idea that comprehension is an incremental process, with listeners building a representation of the sentence as each word is heard. For comprehension to be efficient, there is an ongoing dynamic consideration of possible interpretations that are constrained by context, semantics, syntax, phonology, and any number of other cues (e.g., MacDonald,

^aThe University of Iowa, Iowa City

^bThe University of Delaware, Newark

^cBoys Town National Research Hospital, Omaha, NE

^dCalifornia State University, Fullerton

Correspondence to Jessica E. Hall: jessicahall@email.arizona.edu

Editor-in-Chief: Sean Redmond

Editor: Lisa Archibald

Received July 14, 2017

Revision received January 18, 2018

Accepted October 7, 2018

https://doi.org/10.1044/2018_JSLHR-L-17-0267

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

Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994). The listener shifts expectations with each additional bit of information because the possibilities for interpretation can either narrow or widen at both the level of phrase (or the “global” level) and the level of the word (or the “local” level). This widening or narrowing of options can be captured in an online task that gauges expectations through eye or hand movements toward pictures depicting possible interpretations of a word or sentence. This type of task is known as the “visual world paradigm.” Typically, stimuli in these tasks are designed to include two or more “competing” interpretations. The amount that participants attend to alternative interpretations can reveal how listeners rely on incoming information during comprehension. In eye tracking tasks, for example, efficient comprehension is evident in patterns of eye fixations toward one target item, whereas more fixations toward alternative interpretations when information is conflicting or ambiguous can indicate less efficient or disrupted comprehension (e.g., see Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In mouse tracking, the same effect is documented via examination of the curvature of the movement trajectory toward the target interpretation (e.g., see Spivey, Grosjean, & Knoblich, 2005). Straight trajectories indicate limited competition, and curved trajectories indicate consideration of the alternative response.

Verbs are an important predictive cue in English and thus have been studied extensively by researchers who seek to understand real-time language processing (e.g., Ferreira & Henderson, 1990; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Holmes, 1984; MacDonald, 1994; Trueswell, Tanenhaus, & Kello, 1993). Essentially, certain verbs are more likely to appear in certain syntactic constructions. The phenomenon of verb bias is the expectation of a particular syntactic construction or interpretation based on distributional properties of the verb. Because the brain is attuned to this probabilistic information, we form biases that drive us to expect the more likely syntactic construction upon hearing the verb. In a child-friendly version of the visual world paradigm, Snedeker and Trueswell (2004) showed that adults and children as young as 5 years old are sensitive to verb bias. In their experiment, the participants heard ambiguous sentences in which a prepositional phrase could be interpreted either as the instrument for carrying out the action or as information modifying the noun to be acted upon. In general, in English, the phrase “with the X” refers to an instrument for executing an action, which we will term the *global bias*. For instance, in the sentence “Poke the camel with the feather,” the tendency is to assume that the feather is being used for poking (instrument interpretation), rather than as a means of describing which camel should be poked (modifier interpretation). The local or word-level bias of the verb *poke* also predicts the phrase will be interpreted this way. Thus, in the ambiguous sentence “Poke the camel with the feather,” the feather is most likely to be interpreted as an instrument rather than a description of which camel should be poked. In contrast, it is rare for the verb *hug* to require an instrument, and so the local bias associated with the individual

verb outweighs the global bias about how prepositional phrases tend to work. Thus, the prepositional phrase in the ambiguous sentence “Hug the cat with the blanket” is interpreted as providing modifying information about the cat. The distributional properties of the verbs *poke* and *hug* narrow potential syntactic contexts and drive expectations of how to interpret the sentence.

Snedeker and Trueswell (2004) tracked the eye movements of the participants to four items while the participants acted out the sentences. Returning to the example, “Poke the camel with the feather,” the participants saw a visual display that featured a target object (a feather), a target animal with a smaller version of the target object (a camel with a tiny feather), a distractor object, and a distractor animal. For trials with instrument-biased verbs such as *poke*, participants were more likely to perform the action using the target object (using the feather to poke the camel) and also looked more often at that target object. They found the inverse for modifier verb trials, with more use of and looks toward the target animal, showing that, in this instance, local biases at the verb level were stronger than the global bias to interpret *with the* phrases as instruments. From this, the researchers concluded that children as young as 5 years old use distributional information specific to the lexical item to inform their interpretation of sentences.

Although *hug* is statistically more likely to occur without an instrument phrase, semantics also likely play a role in how the above example sentences are interpreted. For example, modifier-biased verbs are often verbs of observation (e.g., *look at*, *listen*, *notice*). Ambridge and colleagues have looked extensively at semantic classes and statistical information in verb argument structure in a series of articles on the retreat from overgeneralization errors in typically developing (TD) children. In these works, the authors find evidence that both semantic and statistical information inform children’s production of errors (see especially Ambridge, Pine, Rowland, & Young, 2008). As Snedeker and Trueswell acknowledge in the discussion of their 2004 study, it is difficult to disentangle the roles of syntactic and semantic information in real language stimuli, when likely this information is learned and used simultaneously.

Psycholinguistic researchers have traditionally employed eye tracking to measure competition dynamics, considering changes in the proportion of looks to the target as revealing patterns of competition between different interpretations. In this case, discrete eye movements (looks) are tallied and recorded. It is also well established that these dynamics can be revealed in hand movements when using a computer mouse. Hand movements reflect the blending of two motor commands yoked to competing cognitive representations (Tipper, Howard, & Houghton, 2000). Thus, mouse movement trajectories represent a continuous measure of dynamic competition as a decision unfolds during processing, as opposed to the discrete movements of the eye. Mouse tracking has been shown to reveal graded spatial attraction by recording the streaming *x,y* coordinates of continuous

computer mouse movements as a participant moves the mouse toward a pictured interpretation of an unfolding linguistic stimulus. The trajectory of the mouse thus illuminates the strength of the competition (Farmer, Cargill, & Spivey, 2007; Spivey, 2007; Spivey et al., 2005; see Freeman, Dale, & Farmer, 2011, for an overview). We illustrate this phenomenon below, first with an example from lexical processing and then with an example from sentence processing.

During a spoken word recognition task, Spivey et al. (2005) found that participants showed greater mouse curvature toward a distractor item when it shared phonological features with the target (candy/candle) than when it did not (jacket/candle), although they almost always selected the correct image. This finding is similar to those in an eye tracking study by Allopenna, Magnuson, and Tanenhaus (1998). Similarly, in a study on the processing of syntactically ambiguous sentences, Farmer et al. (2007) compared mouse curvature for temporarily ambiguous sentences such as “Put the apple on the towel in the box” and unambiguous sentences such as “Put the apple *that’s* on the towel in the box.” They demonstrated that, when participants were instructed to “Put the apple on the towel in the box,” their average movement trajectory curved significantly more toward the picture of the alternative (and incorrect) destination (an empty towel) before placing the apple at the correct destination (a box) than the trajectory produced during the unambiguous control sentences (e.g., “Put the apple *that’s* on the towel on the box”). Thus, the overall trajectory indexes the motor commands corresponding to each alternative. Even though participants click on the object corresponding to the ultimately correct interpretation of a temporary ambiguity, the degree of curvature toward the location consistent with the incorrect interpretation reflects, on a trial-by-trial basis, the amount of competition from (or attraction toward) the competing alternative. This was a replication of the findings of the classic eye tracking experiment by Tanenhaus et al. (1995). The curvature difference in each of these studies indexes the amount of competition participants experienced when multiple possibilities were temporarily available as they processed temporally unfolding linguistic information. We note that the similarities between the use of mouse trajectories and eye movements for examining online processing are well established in the literature at this point.

Language Processing in DLD

Children with DLD, also termed *specific language impairment*, show difficulties with comprehension ranging from slower recognition of words (Clarke & Leonard, 1996) to greater difficulty in understanding sentences that are long (Montgomery, 2000), syntactically complex (Montgomery & Evans, 2009), or semantically loaded (Leonard, Deevy, Fey, & Bredin-Oja, 2013) to challenges understanding complex *wh*-questions (Deevy & Leonard, 2004). Fewer studies of comprehension in adults with DLD exist, but several lines of evidence point to comprehension problems persisting beyond childhood. A review by Law, Garrett,

and Nye (2004) suggests that the few receptive language interventions that have been tested are ineffective, and longitudinal studies show evidence of sustained difficulties on language and academic measures into adulthood (e.g., Dockrell, Lindsay, & Palikara, 2011; Durkin & Conti-Ramsden, 2007). These findings, taken together, are evidence that the comprehension difficulties that characterize DLD are resistant to treatment and do not disappear due to maturation.

The advantage of sentence processing tasks is that they can measure the process, in addition to measuring the outcome of comprehension. Understanding the process could be revealing for understanding comprehension breakdown for individuals with DLD. Few studies have examined language processing in DLD, and to our knowledge, only one had adult participants (Poll, Watkins, & Miller, 2014). These studies display a variety of differences between individuals with DLD and their typical peers: in the amount of eye movements to the target (Andreu, Sanz-Torrent, & Trueswell, 2013) or to distractors (Borovsky, Burns, Elman, & Evans, 2013; McMurray, Munson, & Tomblin, 2014; McMurray, Samelson, Lee, & Tomblin, 2010) and in reaction times (Marinis & van der Lely, 2007; Montgomery, 2005; Norbury, 2005; Poll et al., 2014). Most of these studies showed poorer word recognition or lexical activation by the DLD group. This is important because, as several of these researcher groups note, word-level deficits could contribute to poorer sentence-level comprehension. More studies that focus specifically on sentence processing in DLD are needed to better understand the connection between word and sentence-level processing. We report on both types of studies here for a more thorough presentation of what is presently known about language processing in DLD.

In her online study of lexical ambiguity, Norbury (2005) found that children and adolescents with DLD were less accurate than their typical peers for the less common meaning of a word, even when verb semantics were a cue. For example, when listening to the sentence “Bill fished from the bank,” children were asked to respond if a picture of money went with the sentence. Even though the verb “fish” gives a clue to the meaning of the ambiguous word “bank,” children with DLD had more difficulty than their TD peers rejecting the picture of money, even when knowledge of word meanings was controlled for. This finding could be taken as evidence that individuals with DLD do not consider alternate lexical meanings and therefore show poorer sentence comprehension. However, difficulty in rejecting the more common meaning of a word could also be interpreted as “greater” consideration of alternative interpretations due to a failure to use disambiguating information from the verb efficiently.

In a visual world word recognition task, McMurray et al. (2010) showed that adolescents with DLD were more likely to continue to consider distractors beyond the point at which the TD groups had stopped actively considering those options. That is, when hearing a word such as *sandal* and faced with visual depictions of a target (sandal), a

competitor with a similar onset (sandwich) or offset (candle), along with a distractor, individuals with DLD looked more at the competitors than their TD peers, even after hearing the entire word. The authors found similar timing for first looks to the target, evidence against a generalized slowing account of DLD. Using a computational model to further understand the results, the authors concluded that deficits in underlying cognitive processes did not allow enough consideration of the target to override consideration of competitors. This study provides further evidence that individuals with DLD may show sentence comprehension deficits because of inefficient lexical processing.

Andreu et al. (2013) found similar results in a visual world paradigm word recognition study with Spanish-speaking children with DLD. In the task, participants saw pictures that represented possible direct objects of the sentence they heard, with one object “favored” by the verb. For example, participants heard “The man closes quickly the door,” (in Spanish) with *door*, *cloud*, *tree*, and *stamp* pictured. Children were instructed only to look at the picture that went with the sentence. Similar to the adolescents with DLD in McMurray et al. (2010), children with DLD looked at the appropriate target as quickly as age-matched controls but did not stop looking at the distractors. In sentences with atypical verb–theme pairs, children with DLD showed fewer looks to the target than their age-matched peers. The authors attribute this finding to children with DLD having weak semantic representations for verbs. In support of this conclusion, TD children and adults with lower vocabulary and reading skills have shown less time looking to the target than participants with higher language skills in other online processing studies (Borovsky, Elman, & Fernald, 2012; Mani & Huettig, 2012; Nation, Marshall, & Altmann, 2003). Because of the heavy lifting that verbs do in assigning meaning to an utterance (i.e., who does what to whom), lexical difficulties with verbs may have a broader overall impact on sentence comprehension in DLD.

Borovsky et al. (2013) tested the influence of verb semantics in a sentence processing a visual world paradigm task that employed eye tracking with adolescent participants with and without a history of DLD. The task was designed such that distractor items related to either the meaning of the subject or the meaning of the verb in the sentence. For example, participants heard “The pirate hides the treasure” and saw *bones*, corresponding with “hide,” and a *ship*, corresponding with “pirate,” as well as *treasure*, the target, and an unrelated item, a *cat*. The objective was to determine if hearing the verb increased consideration for the verb-related competitor. The group with DLD did not differ from their TD peers in the number of looks or how quickly they looked at the target. However, participants with DLD looked less to the verb-related distractor item after hearing the verb than their TD peers, evidence for verb-related difficulty. In contrast to the conclusions of McMurray et al. (2010), the authors speculate that processing difficulties limited consideration of related or possible meanings, which provides an explanation for how the DLD group could respond as quickly

as the TD group and show fewer looks to distractors. In this case, the participants with DLD could have only entertained one possible sentence meaning, whereas their TD peers actively considered several. Note that the differences found between this study and that of McMurray et al. (2010) may be due to differences in processing an entire sentence versus a one word only.

Finally, Poll et al. (2014) examined reaction time in a word detection task with adults with and without DLD. They found slower reaction times for the DLD group than the TD group, which, they concluded, may have been due to an inability to suppress interfering information; that is, they paid too much attention to the sentence rather than focusing on listening to the target word. The DLD group was slightly faster on trials with a slow rate of presentation than trials with a normal presentation speed, which the authors interpreted as indicating differences between groups in dynamic mechanisms supporting working memory. These results follow a similar pattern to a word detection task with children with DLD (Montgomery, 2005). Furthermore, a study with TD children showed a relationship between sentence processing and working memory (King & Just, 1991).

Sentence processing and lexical competition models often disregard contributions of general cognitive mechanisms that are more consistent with an information-processing approach to language learning and use. However, the information-processing approach is integral to accounts of DLD. Thus, we consider these issues briefly. Because timing has been shown to be similar between diagnostic groups for at least some types of language tasks (Andreu et al., 2013; Borovsky et al., 2013; McMurray et al., 2010), we do not consider generalized slowing to be the main factor in processing difficulties in DLD. Instead, we posit that deficits in cognitive mechanisms that support working memory and attention could limit the amount of information that can be taken in or utilized during sentence processing by individuals with DLD (Ellis Weismer, Evans, & Hesketh, 1999; Gillam, Cowan, & Marler, 1998; Hoffman & Gillam, 2004; Leonard et al., 2007; Montgomery, 2000; Montgomery & Evans, 2009). Alternatively, McMurray et al. (2010) and Poll et al. (2014) propose that participants with DLD had difficulty in suppressing competing information in their online language processing studies. In either case, deficits in how efficiently verb bias information is learned or utilized during language processing could result in comprehension deficits.

Lexically based information such as verb bias may be difficult for individuals with DLD to hold in memory during a long or complex sentence, rendering verb bias information unusable to the listener. If verb bias information is not usable to make predictions, participants with DLD may instead use a compensatory strategy (e.g., rote memorization or strong association with the most common possibility) that decreases competition by limiting consideration of competitors that typical individuals would consider viable. This may be an effective strategy some of the time, but it would also lead individuals with DLD to disregard

interpretations that should be considered. For instance, one could imagine that individuals with DLD could rely on global biases as a compensatory strategy if local biases are not available to them. These global biases could lead to less consideration of competitors, as well as responses that are inconsistent with verb bias or even incorrect responses.

Objective

In this study, we explored the idea that individuals with DLD might experience language processing difficulties, consistent with the well-documented comprehension deficits in this population. We adapted materials from Snedeker and Trueswell (2004) because a task with ambiguous sentences may provide a more sensitive test of differences than those with simple unambiguous sentences. Because their task was successfully performed by both adults and 5-year-olds, it should be accessible to adults with impaired language. Our research question asks whether individuals with DLD use verb bias in their interpretation of ambiguous sentences. We hypothesized that, if individuals with DLD do not use verb bias information, then they will show differences in real-time responses to language as compared to their typical peers.

The role of verb bias is what is under consideration here. Differences in use of information about the verb's likelihood of occurring with modifier or instrument prepositional phrases could disrupt sentence comprehension, changing the manner of the interpretation. The null hypothesis is that adults with DLD use verb bias information efficiently. In this case, we expect that, for both the DLD and TD groups, mouse trajectories would be curved when choice of interpretation does not match verb bias because there will be greater competition from the more likely choice. When the chosen interpretation matches verb bias, mouse trajectories would be straight because the less likely interpretation will pose less competition. The alternative hypothesis is that adults with DLD do not use verb bias information efficiently. In this case, we could see either decreased (straighter trajectories) or increased (more curved trajectories) consideration of alternative interpretations or altogether different patterns of response as compared to their unaffected peers. Decreased consideration would cause mouse trajectories to be straight across all trial types because participants would use something else (such as semantic information or global bias) to resolve ambiguity or possibly not even realize two interpretations are possible. Increased consideration would cause mouse trajectories to be more curved across all trial types because participants would not have the information to determine which is more likely and thus show more confusion or wavering in their choice.

Method

Participants

Participants included 18 adults ages 18–26 years with typical development and 18 adults identified as having DLD,

who were matched with the TD adults on age, gender, and college enrollment. One participant from the DLD group also participated in the study of McGregor (2017), and all others participated in the studies of Hall, Owen Van Horne, McGregor, and Farmer (2017, 2018) and McGregor, Arbisi-Kelm, and Eden (2017). Participants provided written informed consent. Participants in the DLD group responded to an advertisement requesting college students who “struggle with spoken or written language.” Group membership was determined via the process described in Fidler, Plante, and Vance (2011), which included composite results of a spelling test and modified token task (Morice & McNicol, 1985). This identification procedure has good sensitivity and specificity for detecting DLD in adulthood. Official diagnosis of language disorder was not required; all but three participants in the DLD group reported receiving a diagnosis of language, reading, or writing impairments. Participants in the TD group reported no history of these or other relevant diagnoses. All participants had hearing within normal limits (American Speech-Language-Hearing Association, 1997), had normal or corrected-to-normal vision by self-report, passed a nonverbal IQ test (Kaufman Brief Intelligence Test–Second Edition, Matrices; Kaufman & Kaufman, 2004), and had no history of autism spectrum disorders per self-report. One TD participant was excluded when he revealed that he was bilingual. In addition, all participants performed the Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007) and the Author Recognition Task (Acheson, Wells, & MacDonald, 2008) to obtain measures of vocabulary ability and text exposure. Participant information is given in Table 1.

In Table 1, one can see a difference of almost a standard deviation between groups on the Kaufman Brief

Table 1. Participant demographic and test means (*SD*) by diagnostic group (typically developing [TD] and developmental language disorder [DLD]), after excluding participants as described in the screening measures.

Measure	TD	DLD	<i>d</i>
<i>n</i>	16	17	
Age	21.0 (1.9)	20.7 (1.1)	0.21
Education (years)	14.1 (1.9)	13.9 (1.1)	0.12
KBIT-2 SS	108.9 (12.3)	99.1 (10.5)	0.85*
Spelling	11.7 (2.5)	4.2 (2.6)	2.96***
Token	40.1 (3.1)	35.2 (5.5)	1.08**
PPVT-4 Raw	206.8 (6.8)	199.4 (11.2)	0.80*
ART	18.9 (7.0)	16.4 (12.7)	0.25

Note. Cohen's *d* represents effect size of group differences for each measure for participants included in analyses. Scores on the Kaufman Brief Intelligence Test–Second Edition, Nonverbal subtest (KBIT-2) are standard scores (SS) with a normative mean of 100 and a standard deviation of 15. Scores on the Spelling and Token tasks are raw counts of items correct out of 15 and 44, respectively. Scores on the Peabody Pictures Vocabulary Test–Fourth Edition (PPVT-4) are raw scores. Scores on the Author Recognition Task (ART) are out of a range of –65 to 65, with 65 being the highest possible score.

p* < .05. *p* < .01. ****p* < .0001.

Intelligence Test–Second Edition. Tomblin and Nippold (2014) acknowledge the tendency for language and nonverbal intelligence scores to correlate but advise against matching diagnostic groups based on nonverbal intelligence or controlling for this factor because of a larger degree of measurement error in children with DLD. We also might expect a wider discrepancy of nonverbal intelligence between groups in a sample of postsecondary students.

Stimuli

The set up for this experiment is a mouse-tracking adaptation of the visual world paradigm, in which two pictures depicting different interpretations of a sentence are shown in opposite corners of a computer screen. We note that the choice of mouse tracking was purely pragmatic and had to do with access to technology and portability of the experimental setup rather than any attempt to validate a methodological approach, given that mouse tracking is well established. Sentences in experimental trials were modified from Snedeker and Trueswell (2004) to be imageable by adding animals in the subject position. Verbs in the 16 sentences were biased to appear with (a) instrumental phrases, for example, “*The elephant pokes the camel with the feather*,” where the verb “poke” is most likely to occur with an instrumental phrase, or (b) modifier phrases, for example, “*The bear chooses the duck with the fork*,” where the verb “choose” is most likely to occur with a modifier phrase. Verb biases were determined from a sentence completion study reported in Appendix B of Snedeker and Trueswell (2004) and a similar sentence completion task reported in Ryskin, Qi, Duff, and Brown-Schmidt (2017) that used a similar verb set. The two possible interpretations were pictured (see Figure 1). We also included each of these sentences with only one possible interpretation pictured, with the alternative picture showing the same two animals and object but not the action indicated by the sentence (i.e., an elephant holding a feather while using its trunk to poke

a camel). These were included to control for the participants’ ability to comprehend the sentences and ensure that they were fully participating in the task. We refer to them as “comprehension trials.” The object nouns in all sentences were chosen from lists provided in Snedeker and Trueswell (2004) for their lack of bias to appear as either an instrument or a modifier. Table 2 lists the verbs used in this experiment, their bias, and their object nouns. Table 3 lists experimental variables and their meanings. See the Appendix for audio and visual stimuli descriptions.

In Snedeker and Trueswell’s (2004) sentence completion study, they report the percentage of time each verb was used with an interpretation associated with the bias that they ultimately use for categorization. For example, the instrument-biased verb “poke” appears with an instrument interpretation 29% of the time, and the modifier-biased verb “hug” appears with a modifier interpretation 19% of the time. These numbers may seem small, but they are large in comparison to the relative proportion of time each verb appeared with other interpretations. For instance, in comparison, “poke” never appeared with a modifier interpretation, whereas “hug” appeared with an instrument only 5% of the time. These norms are from 2004, but Ryskin and colleagues ran the same norming experiment again in their 2017 study and confirmed the categorization of nine of the 16 verbs that were also in Snedeker and Trueswell (2004). Ryskin et al. (2017) excluded the seven remaining verbs because they were not amenable to the touch screen that they used in their study. We report strength values from Snedeker and Trueswell (2004) and Ryskin et al. for each verb in Table 2.

Eight practice trials familiarized participants with the task using sentences and pictures distinguished by color, shape, and size information (e.g., “The green triangle is bigger than the black triangle”). Twenty-four control trials that involved no competition of interpretation (e.g., “click on the dolphin that is smiling,” in which picture choices are a dolphin and a rooster) were interleaved randomly

Figure 1. On an experimental trial, the participant hears a sentence such as “The elephant pokes the camel with the feather” and sees two possible interpretations: the instrument interpretation, on the left, consistent with the bias of the verb *poke*, and the modifier interpretation, on the right, inconsistent with verb bias. On comprehension trials, only one of these two interpretations is shown, and the other picture is an impossible interpretation.

Table 2. List of verbs by type and the nouns that appear with the verbs in the sentences.

Bias	Verb	Strength (Snedeker & Trueswell, 2004)	Strength (Ryskin et al., 2017)	Instrument/modifier
Instrument	hit	24	92	flower
	tickle	24		fan
	poke	29	88	feather
	clean*	38	86	t-shirt
	bop*	38	87	ball
	brush*	43		sponge
	cover*	43		book
	feed*	43	71	glass
	look at	19	65	glass
	hug	19	77	blanket
Modifier	find	24	72	stick
	talk to	24		tube
	sing to	29		funnel
	yell at*	30		funnel
	listen to*	38		tube
	choose*	81	94	fork

Note. Strength of bias is determined by the percentage of time that a verb appeared with the biased interpretation in the norming sentence completion study reported in Appendix B of Snedeker and Trueswell (2004). Verbs with asterisks denote “strongly biased” verbs.

with experimental trials. These trials were included to control for motor abilities because DLD is often comorbid with motor impairment (Hill, 2001).

Language Processing Measure

Using MouseTracker mouse-tracking software (Freeman & Ambady, 2010), we recorded the path the mouse took on each trial toward the target location, using *x,y* coordinates to determine curvature. From those recordings of the path, we used maximum deviation to represent mouse curvature. This measure is calculated as the farthest point in the actual movement trajectory from an idealized straight line from the starting position to the ending position. It reflects the maximum amount of competition experienced by a participant toward the nonselected response on a trial-by-trial basis. Thus, we have one value for each trial to indicate the amount of competition. As calculated by the software, high maximum deviation values indicate trajectories that curved toward the nonselected response, and negative values indicate trajectories that curved away from the nonselected response. Figure 2 provides

an example illustration. The solid lines demonstrate averaged mouse trajectories for each diagnostic group and condition, and the dashed line is the idealized straight line drawn from the starting point to the end point of a trajectory, from which the maximum deviation is measured. Although the mouse trajectory is a real-time measure, maximum deviation does not capture time-based information. Thus, it is a measure of overall competition without respect to time course dynamics. Average maximum deviation appears to index competition across a number of types of stimuli (auditory lexical decision: Krestar, Incera, & McLennan, 2013; covert memory strength: Papesch & Goldinger, 2012; evaluative thinking: McKinstry, Dale, & Spivey, 2008).

The images displayed in this task are visually complex, each involving one to two “actors,” an instrument, and an action. Although participants did have time to view the images before listening to the sentence, we reasoned that language comprehension might be negatively impacted given this complexity should participants attempt to move while listening. Thus, we required participants to listen to the entire sentence before beginning their movements. Movements were not recorded while the audio was playing to

Table 3. Task variables and their meanings.

Variable type	Variable	Definition	Values
Dependent variable	Maximum deviation	The measure of competition from the interpretation not chosen, calculated as the farthest point in the actual trajectory from an idealized straight line from start to end points	0 represents straight trajectories and little competition; positive values represent curved trajectories and more competition
Independent variables	Bias	The interpretation favored by statistics	Modifier/instrument
	Consistency	The interpretation chosen by the participant compared to verb bias	Consistent/inconsistent
	Strength	The expected inherent strength of the verb bias	0–100, with 0 representing weak bias and 100 representing strong bias

Figure 2. Average mouse trajectories for choosing instrument (left) versus choosing modifier (right) on instrument-biased trials, with the group with developmental language disorder (DLD) shown in gray and the group with typical development (TD) shown in black. The dashed line shows the ideal straight line from which maximum deviation for each trajectory is measured.

preserve the comprehension component of the task and to reduce task interference effects across diagnostic groups. The degree to which this decision may have reduced putatively “early” effects of ambiguity on trajectories remains an empirical question. It is likely, however, that allowing earlier movement would only exacerbate findings of group differences and further confound processing difficulties with task demands. In addition, as a result of this decision, we focus our analyses on maximum deviation values and not on more derivative measures, such as acceleration rate and time-normalized trajectories.

Procedure

Procedures were approved by the institutional review board at the University of Iowa. Adults were invited to participate in research activities during a 1-hr visit. Participants completed the Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007) and the Author Recognition Task (Acheson et al., 2008) prior to doing the experimental task. At the start of the task, instructions directed participants to use a computer mouse to choose the best pictured interpretation of each sentence that they hear. In each trial, two pictures were displayed in the top left and right corners of the screen. Pictures were counterbalanced for where the instrument interpretation appeared and where the biased interpretation appeared (experimental and comprehension trials), and where the correct interpretation appeared (control and comprehension trials). Participants were instructed to take their time considering the pictures and to press the button to hear the sentence only when they were sure they understood how

the pictures differed. They could look at the pictures for as long as they liked. After the sentence ended, a star appeared, which was the sign to move the mouse as quickly as possible to their chosen interpretation of the sentence. Mouse movements were only recorded once the star appeared. If the participants took longer than 3 s to initiate mouse movement, a warning message was displayed after the trial encouraging them to try to choose as quickly as possible. Participants first completed the eight practice trials and then the 16 experimental trials, 16 comprehension trials, and 24 control trials, which were interspersed, and the order was randomized for each participant.

Analysis

We used the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) and the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) in R Version 3.4.3 (R Core Team, 2017) to run linear mixed-effects models to explore all comparisons of interest. The Akaike information criterion was used to determine the best fit model, using a maximal random effects structure if the difference in Akaike information criterion between models was less than 2. See Table 3 for a summary of variables.

Comprehension and Choice of Interpretation

We first analyzed differences in accuracy for comprehension trials between groups through a two-sample *t* test. This was also used as a screening measure; participants with less than 70% accuracy were excluded. Because the pictures in the comprehension trials should clearly direct

interpretation and participants are adults, we did not expect differences in accuracy.

For experimental trials, we used mixed-effects logistic regression to examine the likelihood that participants would choose an instrument interpretation. We used this type of model because the variable of interest was a measure of likelihood of choice between two possibilities. We included the bias of the verb (instrument or modifier) and diagnostic group and their interaction as predictors to determine whether individuals with DLD have different knowledge of verb bias from TD individuals. Instrument bias and the TD group serve as the reference categories, and variables were effects-coded.

We expected all participants to choose instrument interpretations slightly more often overall because of a sensitivity to the global bias because this is the more common interpretation in English. We expected that individuals with DLD would show different response patterns from their TD peers if they did not use local information to resolve the ambiguity. If they did use this information, patterns might be similar across groups for choice of interpretation, and differences might only be detected in mouse trajectories.

Mouse Curvature

The measure used to determine sensitivity to verb bias in online processing was the amount of mouse curvature, measured in maximum deviation. We included diagnostic group (TD or DLD), the bias of the verb (instrument or modifier), and whether the participant's choice was consistent with bias (consistent or inconsistent) and their interaction as fixed effects in the model. Average maximum deviation for control trials was included to control for differences in motor control. For all analyses, instrument bias, consistent choice, and the TD group served as the reference categories, and variables were dummy coded.

We expected that TD participants would show sensitivity to verb bias and would thus show straighter trajectories on trials when their choice was consistent with the bias of the verb (i.e., when they select the instrument interpretation when the sentence contains an instrument-biased verb) than on trials when their choice was inconsistent. We expected TD participants would show more curved trajectories, as represented by maximum deviation, when their choice was inconsistent with bias (i.e., when they select a modifier interpretation for a trial with an instrument-biased verb) because they would consider the alternative more, and thus, their mouse movement would reflect that. A main effect or interaction with the bias of the verb that showed increased curvature on modifier-biased trials would indicate sensitivity to the global instrument bias.

We expected adults with DLD to be less sensitive to verb bias information than their TD peers and thus show different mouse trajectory patterns. It is possible that they could show less consideration of the alternative on all trials (and thus have straight trajectories) because weaker sensitivity could force them to choose at random.

We include average maximum deviation on control trials as a measure of baseline motoric variation for each participant. However, this does not account for variability of the bias that might be associated with the verb. We account for this by using a measure of expected inherent strength of verb bias from the norming data provided by the original Snedeker and Trueswell (2004) study described previously and shown in Table 2. We did not use the more recent values from Ryskin et al. (2017) because of the missing values from the verbs they did not use in their study. We use these percentages as a covariate measure in our model that we characterize as strength of bias. By using a measure for the verb, we isolate variability associated with the linguistic stimuli separate from variability that may arise from the visual stimuli. We expect the strength of bias to interact with consistency in predicting the amount of curvature in the trajectories. Thus, we predict larger maximum deviations for strongly biased verbs when participants choose inconsistent interpretations than more weakly biased verbs or when participants choose consistent interpretations. Strength may also interact with diagnostic group if TD participants show more sensitivity to verb bias than participants with DLD.

Results

Data Screening (Comprehension and Trajectories)

Practice trials were not included in any analysis. As mentioned previously, we measured accuracy on comprehension trials to detect differences between groups. We excluded participants who answered more than four of the 16 comprehension trials incorrectly. One participant in the DLD group and one participant in the TD group were excluded for low accuracy on the comprehension trials. We reverified that group differences in participant characteristics held after we excluded these two participants. Group means and Cohen's *d* reported in Table 1 reflect these exclusions. On average, each participant in the DLD group made 1.2 ($SD = 1.5$) incorrect responses (choosing a not possible interpretation in an unambiguous context), and each participant in the TD group made 0.8 ($SD = 1.0$) incorrect responses for the 16 comprehension trials. A Mann-Whitney *U* test confirmed that groups did not significantly differ in the number of incorrect responses, $U = 121$, $p = .60$.

For experimental trials of retained participants, we screened every trajectory individually for aberrant movements (i.e., noninterpretable looping cycling leftward and rightward; Freeman, Ambady, Rule, & Johnson, 2008; 28 trials in each group and 56 trials in total). We then excluded trials with too long a reaction time ($> 5,000$ ms; one trial in total) and trials in which the participant took more than 2 s to initiate movement (two trials in total). Overall, 11% of experimental trials were discarded. We did not exclude any participants for missing trials because mixed-effects models can adequately handle missing data, and we analyzed individual trial-level data. The largest numbers of

trials missing for a participant were nine in the TD group and eight in the DLD group out of a total of 16 experimental trials. Visual inspection of trajectories showed similar movements for left and right responses, and thus, we did not include response position as a covariate in our models.

Choice of Interpretation

We first consider group differences in the choice of interpretation on experimental trials as indication of differences in comprehension in DLD. In the mixed-effects logistic regression, the best fit model included random intercepts for subject and item. Table 4 reports log odds (reported as *Beta*), and Figure 3 shows the mean percentage of time participants in each diagnostic group chose instrument interpretations for trials with instrument- and modifier-biased verbs. Consistent with the global bias of “with the X” prepositional phrases to be interpreted as instruments, both groups chose instrument interpretations more often than modifier interpretations, regardless of verb bias. The probability of selecting an instrument interpretation across both instrument and modifier verb types and across groups was 77%, $z = 3.15$, $p = .002$. There was a significant effect of verb bias, such that when the verb had an instrument bias, participants across groups were 89% likely to choose an instrument interpretation, compared to 59% when the verb had a modifier bias, $z = 2.46$, $p = .01$. There was no main effect of group, $p = .95$, but group moderated the verb bias effect, $z = 2.52$, $p = .01$. The probability values of TD participants choosing an instrument interpretation were 92% with an instrument-biased verb and 51% with a modifier-biased verb. In comparison, DLD participants had 86% probability of choosing an instrument interpretation with an instrument-biased verb and 66% with a modifier-biased verb. Although both groups showed sensitivity to both global instrument bias and lexically based biases, the TD group appears more sensitive to the bias of the verb in their choice of interpretation.

Table 4. Summary of logistic regression analysis for variables predicting choice of instrument interpretation.

Factor	Variance	SD	β	SE	p
Random factors					
Subject	0.77	0.88			
Item	1.95	1.40			
Fixed factors					
(Intercept)			1.23	0.39	.002
Bias (reference category: instrument)			0.88	0.36	.01
Group (reference category: TD)			−0.01	0.19	.95
Bias × Group			0.31	0.12	.01
Model $\chi^2 = 8.26$					

Note. TD = typically developing.

Figure 3. Mean percentage of trials by group (typically developing [TD] and developmental language disorder [DLD]) of choice of interpretation (instrument, modifier) for instrument- and modifier-biased trials. Error bars represent standard error.

Mouse Curvature

We ran a model with only control trials included to test for baseline group differences. Groups did not significantly differ on maximum deviation on control trials, $p = .28$, as can be observed in Figure 4. The negative means indicate that the trajectories for both groups, on average, curved away from the unselected response, indicating very little competition. This is expected on completely unambiguous control trials.

We now consider the verb bias effect as evidenced by mouse curvature in a mixed-effects linear regression model. Greater mouse curvature on trials in which participants' choice was inconsistent with verb bias than when choice was consistent with verb bias indicates sensitivity to local verb bias. Greater mouse curvature on modifier-biased verb trials than on instrument-biased verb trials indicates sensitivity to the global instrument bias. We used a linear mixed-effects model with the following fixed effects: bias of the verb (instrument or modifier), whether the participant chose an interpretation that was consistent with bias (consistent or inconsistent), the expected strength of verb bias (0–100), diagnostic group (TD or DLD), and their interaction, as well as average maximum deviation on control trials. Table 5 reports model results. Figure 5 displays the influence of verb bias and consistency of choice of interpretation on mouse curvature for each diagnostic group, and Figure 6 displays the influence of these factors as they interact with strength of verb bias. The y -axis represents maximum deviation; thus, positive values represent deviation toward the unselected response, and values near zero represent straight trajectories. Mean maximum deviation on control trials did not significantly contribute, $p = .72$. We saw a main effect of verb bias, $t(435.3) = 2.42$, $p = .02$. There were two-way interactions between consistency of choice and diagnostic group, $t(467.1) = 2.90$, $p = .004$, and consistency of choice and verb bias strength, $t(467.3) = 2.11$, $p = .04$. There was a three-way interaction between verb bias, consistency of choice, and diagnostic

Figure 4. Diagnostic group (typically developing [TD] and developmental language disorder [DLD]) means for maximum deviation on control trials.

group, $t(465.8) = -2.46$, $p = .01$, as can be observed in Figure 5, with a similar interaction between consistency and bias for both groups, but overall less deviation in the DLD group. There was also a three-way interaction between consistency of choice, diagnostic group, and verb

bias strength, $t(466.7) = -2.98$, $p = .003$. Finally, the four-way interaction between these variables was significant, $t(464.1) = 2.54$, $p = .01$. As Figure 6 shows, the TD group demonstrates the expected pattern for interaction between consistency and strength of verb bias for

Table 5. Results of the model showing the influence of consistency of choice of interpretation, verb bias, expected strength of verb bias, and their interaction, as well as average maximum deviation on control trials on maximum deviation.

Factor	Variance	SD	β	SE	p
Random factors					
Subject	0.00	0.01			
Fixed factors					
(Intercept)			-0.18	0.19	.35
Consistency of choice of interpretation (reference category: consistent)			-1.45	0.80	.07
Verb bias (reference category: instrument)			0.53	0.22	.02
Diagnostic group (reference category: TD)			-0.02	0.27	.93
Strength of verb bias			-0.006	0.005	.30
Verb Bias $\times$ Consistency of Choice			1.09	0.08	.18
Verb Bias $\times$ Diagnostic Group			-0.16	0.31	.60
Verb Bias $\times$ Strength of Verb Bias			-0.009	0.006	.12
Consistency of Choice $\times$ Diagnostic Group			2.58	0.89	.004
Consistency of Choice $\times$ Strength of Verb Bias			0.04	0.02	.04
Diagnostic Group $\times$ Strength of Verb Bias			0.0004	0.008	.96
Verb Bias $\times$ Consistency of Choice $\times$ Diagnostic Group			-2.27	0.92	.01
Verb Bias $\times$ Consistency of Choice $\times$ Strength of Verb Bias			-0.03	0.02	.08
Verb Bias $\times$ Diagnostic Group $\times$ Strength of Verb Bias			0.003	0.008	.74
Consistency of Choice $\times$ Diagnostic Group $\times$ Strength of Verb Bias			-0.07	0.02	.003
Verb Bias $\times$ Consistency of Choice $\times$ Diagnostic Group $\times$ Strength of Verb Bias			0.06	0.02	.01
Average maximum deviation on control trials			0.10	0.26	.72

Note. TD = typically developing.

Figure 5. Mean maximum deviations of mouse trajectories between groups (typically developing [TD] and developmental language disorder [DLD]) by verb bias (instrument and modifier) and consistency of choice (consistent and inconsistent). Error bars represent standard error.

Figure 6. Mouse trajectories are impacted by the expected strength of verb bias, with higher values representing more strongly biased verbs, and the consistency of choice of interpretation for the typically developing (TD) group but not the group with developmental language disorder (DLD).

instrument-biased trials, with higher maximum deviation for more strongly biased verbs when their choice was inconsistent with the bias than for weakly biased verbs. The DLD group does not show this pattern, and we also do not see this pattern for either group on modifier-biased trials. Figures 5 and 6 suggest that the DLD group had straighter trajectories than the TD group when choosing modifier interpretations across all trials. Although there was no theoretical reason to test this difference, it suggests less sensitivity to the global instrument bias than the TD group.

Initiation Time

Given these results, we decided to compare initiation time between groups to determine if participants with DLD were simply waiting longer than the TD group to move the mouse, thus masking competition effects. We used a linear mixed-effects model with initiation time as the dependent variable, with fixed effects of the bias of the verb (instrument or modifier), whether the participant chose an interpretation that was consistent with bias (consistent or inconsistent), the expected strength of verb bias (0–100), diagnostic group (TD or DLD), and their interactions. A random subject intercept was supported by the data. Table 6 reports model results, which are illustrated in Figures 7 and 8. There was no main effect of group, $p = .97$, and no interaction between group and bias, $p = .97$, group and consistency, $p = .11$, or group and strength of verb bias, $p = .98$, and three-way interactions with group were also not significant, $ps > .05$, indicating that participants with DLD are taking the same amount of time to make their choice of interpretation as the TD group, even when choosing interpretations that are inconsistent with the bias of the verb. The

four-way interaction was marginal, $t(441.4) = -1.91$, $p = .057$. Figures 7 and 8 also demonstrate that the DLD group was as fast or faster than the TD group, and so we do not see evidence that they are waiting longer to initiate movements, masking competition.

Discussion

Efficient comprehension takes place through dynamic consideration of competing meanings that are constrained by any number of cues in the environment. Here, we designed a visual world paradigm experiment to test sensitivity to one such cue, verb bias. We used this experimental paradigm to examine the process of comprehension in adults with and without DLD.

As we predicted, when interpreting syntactically ambiguous sentences in the visual world, adults with DLD showed less sensitivity to verb bias than their TD peers in the choice of interpretation. Specifically, they chose an interpretation that matched the bias of the verb less often than the TD participants. Furthermore, by tracking computer mouse movements as participants chose an interpretation, we found that TD participants' trajectories showed more evidence of sensitivity to verb bias. They displayed more curved trajectories when their choice did not match bias than when it matched for instrument-biased trials than the participants with DLD. We tested the influence of strength of verb bias and found differences between groups in how it affected mouse trajectories. Visual inspection of maximum deviation values suggests overall straighter trajectories for the DLD group when choosing modifier interpretations. We will consider how these results fit with previous findings on online processing in DLD, but first, we address a limitation in the interpretation of findings.

Table 6. Results of the model showing the influence of consistency of choice of interpretation, verb bias, expected strength of verb bias, and their interaction, as well as average maximum deviation on control trials on initiation time.

Factor	Variance	SD	β	SE	p
Random factors					
Subject	31,074	176.3			
Fixed factors					
(Intercept)			90.59	130.49	.49
Consistency of choice of interpretation (reference category: consistent)			683.72	528.06	.20
Verb bias (reference category: instrument)			237.14	140.50	.09
Diagnostic group (reference category: TD)			-7.45	187.97	.97
Strength of verb bias			3.24	3.47	.35
Verb Bias $\times$ Consistency of Choice			-708.99	539.50	.19
Verb Bias $\times$ Diagnostic Group			-8.16	203.89	.97
Verb Bias $\times$ Strength of Verb Bias			-6.31	3.77	.10
Consistency of Choice $\times$ Diagnostic Group			-941.66	589.14	.11
Consistency of Choice $\times$ Strength of Verb Bias			-14.63	13.37	.27
Diagnostic Group $\times$ Strength of Verb Bias			0.13	5.00	.98
Verb Bias $\times$ Consistency of Choice $\times$ Diagnostic Group			949.62	609.66	.12
Verb Bias $\times$ Consistency of Choice $\times$ Strength of Verb Bias			17.93	13.74	.19
Verb bias $\times$ Diagnostic Group $\times$ Strength of Verb Bias			0.58	5.45	.92
Consistency of Choice $\times$ Diagnostic Group $\times$ Strength of Verb Bias			25.58	15.15	.09
Verb Bias $\times$ Consistency of Choice $\times$ Diagnostic Group $\times$ Strength of Verb Bias			-30.05	15.76	.057

Note. TD = typically developing.

Figure 7. Mean initiation times for both diagnostic groups (typically developing [TD] and developmental language disorder [DLD]) by verb bias (instrument and modifier) and consistency of choice (consistent and inconsistent).

Figure 8. Initiation time is not impacted by the expected strength of verb bias for either diagnostic group (typically developing [TD] and developmental language disorder [DLD]). Shading represents standard error of the model.

Limitations

It is possible that participants may have used the semantic information in the sentence to disambiguate. For example, they may have considered whether the object in the sentence could plausibly be used as an instrument or whether a certain animal was likely to perform the action with or without an instrument (e.g., an animal with fins instead of paws). We attempted to control for this by using normed direct objects from the original study (Snedeker & Trueswell, 2004), which had similar plausibility for use as instruments or modifiers. However, if adults with DLD could not access verb bias information efficiently, they may have relied more on semantic or contextual cues to resolve ambiguity, as suggested by Borovsky et al. (2013). The salience of objects in the pictures may also have contributed to how much participants considered instrument interpretations. Semantic or visual information likely interfered on modifier-biased trials given that TD adults in our study chose instrument interpretations more often on these trials than in the modifier-biased trials of the original study (Snedeker & Trueswell, 2004). A more recent study with a similar verb set (Ryskin et al., 2017) has shown norming and experimental results that were consistent with Snedeker and Trueswell (2004), and therefore, we do not think it likely that the biases of the verbs in this experiment were problematic. Because our study shows that the TD group is sensitive to verb bias at least for instrument-biased trials, we will not pursue this point further, but we acknowledge that our stimuli limit the conclusions we can draw, and future studies should include normed stimuli when possible.

Sensitivity to Verb Bias Information in DLD

The DLD group showed different patterns of choice and competition than the TD group. Because the DLD group chose the interpretation predicted by verb bias less often than the TD group and showed less evidence of verb bias in their mouse trajectories, it is unlikely they were using verb bias information to resolve ambiguities. The straighter trajectories by this group indicate little competition from the alternative interpretation. It is possible that deficits in cognitive mechanisms supporting working memory limited the ability of individuals with DLD to actively consider alternatives flexibly in the moment. Participants with DLD in some instances may have been unaware that a modifier interpretation was possible and did not recognize the sentences as ambiguous. Although perhaps the global bias influenced their performance, it is unlikely that participants with DLD used global bias alone as they chose modifier interpretations as often as the TD group did across verb types. It is also possible that participants with DLD have verb biases that are idiosyncratic and that do not reflect the verb biases of the general population. Hence, their data would not show the same influence of verb bias strength as predicted by norming data as the TD group. Under this hypothesis, each individual with DLD would have had particular lexical items stored with particular interpretations

and would not have entertained other possibilities for those verbs. Repeating the same lexical item across multiple trials while manipulating context and semantic plausibility would help us to understand the comprehension strategies employed and to determine how global and lexical biases interact.

The patterns of results seen in our study are consistent with findings from other real-time processing studies of DLD. Like the finding by Borovsky et al. (2013) of fewer looks to the action-related distractor by the DLD group, we saw less consideration of the instrument interpretation when choosing modifier by participants with DLD, in contrast to the TD group. We saw similar time course results across groups, consistent with findings by Andreu et al. (2013) and Borovsky et al. (2013). The pattern of similar initiation times but straighter trajectories than the TD group may indicate the DLD group used a compensatory strategy rather than verb bias to resolve ambiguity or even that these sentences were not ambiguous to them. Even if they were aware of an alternative interpretation, adults with DLD in this study may not have been able to actively consider it in the moment, similar to the children with DLD who had trouble recognizing less common word meanings in Norbury (2005). Our findings diverge from those that show increased consideration of alternatives by participants with DLD compared with their TD peers (Andreu et al., 2013; McMurray et al., 2010), though it should be noted that these studies examined word recognition rather than sentence comprehension. To test whether poorer working memory could have dissimilar effects for word and sentence processing, one would need to test the same participants across the two types of tasks. Eye tracking and mouse tracking could also be used in combination to measure competition over the course of the sentence to gauge the impact of memory and attention processes (i.e., do looking and trajectory patterns differ directly after the verb compared with after the sentence has played in its entirety).

Implications and Conclusions

Findings from this study provide support for future work employing real-time methodologies to study language in the DLD population. We saw group differences in their choice of interpretation, but the mouse curvature measures were more revealing for understanding how participants with DLD were comprehending the sentences and whether they utilized verb bias information. The mouse trajectories by adults with DLD suggest that they did not use verb bias information in the same way as their TD peers, if at all. Future studies might examine how context (e.g., the number of competitors present, the presence or salience of visual cues) or other factors such as semantic knowledge affect sentence processing to explore what compensatory strategies the DLD group may have used when interpreting the sentences in this study. It would also be important to determine how word- and sentence-level processing deficits

co-occur and interact within participants in their impact on comprehension.

This study examined sensitivity to verb bias information to resolve ambiguous sentences in adults with DLD and their TD peers. Ambiguous sentence contexts can help us understand how individuals weigh informational cues during comprehension. The primary finding is that adults with DLD showed less sensitivity to verb bias information than their typical peers during sentence processing, which may be related to overall comprehension and proficiency with language. Results are in line with previous findings that suggest limitations in mechanisms that support working memory could affect the ability of individuals with DLD to entertain multiple interpretations during processing. Adults with DLD also made different choices of interpretation than adults with typical development, suggesting that these processing differences may lead to costs in comprehension as well. This study advances our knowledge of DLD in adulthood and is a step toward understanding how maturation and experience with language shape language processing in this population.

Acknowledgments

National Institute on Deafness and Other Communication Disorders Grant 5R01DC011742 supported participant recruitment and the efforts of the third author. The first author's time was supported by National Institute on Deafness and Other Communication Disorders Grant 1F31DC015370. We thank Nichole Eden, Tim Arbisi-Kelm, Danielle Reese, and Dan Plebanek for their assistance with data collection; Genna Valvick and Callum Duff for their help with stimuli creation; and Bob McMurray, Gerry Altmann, and Sarah Brown-Schmidt for their input on data analysis. We also thank the members of Karla McGregor's Spring 2017 Scientific Writing Seminar for extensive feedback on early drafts of this article.

References

- Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, 40(1), 278–289.
- Alloppenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38(4), 419–439.
- Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition*, 106(1), 87–129.
- American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening*. Retrieved from <http://www.asha.org/policy/GL1997-00199/>
- 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.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- 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.
- 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.
- Clarke, M. G., & Leonard, L. B. (1996). Lexical comprehension and grammatical deficits in children with specific language impairment. *Journal of Communication Disorders*, 29(2), 95–105.
- Deevy, P., & Leonard, L. B. (2004). The comprehension of wh-questions in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 47(4), 802–815.
- Dockrell, J. E., Lindsay, G., & Palikara, O. (2011). Explaining the academic achievement at school leaving for pupils with a history of language impairment: Previous academic achievement and literacy skills. *Child Language Teaching and Therapy*, 27(2), 223–237.
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test—Fourth Edition (PPVT-4)*. Minneapolis, MN: Pearson Assessments.
- Durkin, K., & Conti-Ramsden, G. (2007). Language, social behavior, and the quality of friendships in adolescents with and without a history of specific language impairment. *Child Development*, 78(5), 1441–1457.
- Ellis Weismer, S., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 42(5), 1249–1260.
- Farmer, T. A., Cargill, S. A., & Spivey, M. J. (2007). Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language*, 57(4), 570–595.
- Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 555–568.
- Fidler, L. J., Plante, E., & Vance, R. (2011). Identification of adults with developmental language impairments. *American Journal of Speech-Language Pathology*, 20(1), 2–13.
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, 42(1), 226–241.
- Freeman, J. B., Ambady, N., Rule, N. O., & Johnson, K. L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. *Journal of Experimental Psychology: General*, 137(4), 673–690.
- Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, 2, 59.
- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37(1), 58–93.
- 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.
- Hall, J., Owen Van Horne, A., McGregor, K. K., & Farmer, T. (2017). Distributional learning in college students with developmental language disorder. *Journal of Speech, Language, and Hearing Research*, 60(11), 3270–3283.

- Hall, J., Owen Van Horne, A., McGregor, K. K., & Farmer, T. (2018). Individual and developmental differences in distributional learning. *Language, Speech, and Hearing Services in Schools*, 49(3S), 694–709.
- Hill, E. L. (2001). Non-specific nature of specific language impairment: A review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, 36(2), 149–171.
- Hoffman, L. M., & Gillam, R. B. (2004). Verbal and spatial information processing constraints in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 47(1), 114–125.
- Holmes, V. M. (1984). Parsing strategies and discourse context. *Journal of Psycholinguistic Research*, 13, 237–257.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test—Second Edition (KBIT-2)*. Circle Pines, MN: AGS.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30(5), 580–602.
- Krestar, M., Incera, S., & McLennan, C. T. (2013). Using mouse tracking to examine the time course of an auditory lexical decision task. *Ohio Psychologist*, 60, 29–32.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Law, J., Garrett, Z., & Nye, C. (2004). The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 47(4), 924–943.
- Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.
- Leonard, L. B., Deevy, P., Fey, M. E., & Bredin-Oja, S. L. (2013). Sentence comprehension in specific language impairment: A task designed to distinguish between cognitive capacity and syntactic complexity. *Journal of Speech, Language, and Hearing Research*, 56(2), 577–589.
- 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.
- MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9(2), 157–201.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution [corrected]. *Psychological Review*, 101(4), 676–703.
- 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 & Performance*, 38(4), 843–847.
- Marinis, T., & van der Lely, H. K. (2007). On-line processing of wh-questions in children with G-SLI and typically developing children. *International Journal of Language & Communication Disorders*, 42(5), 557–582.
- McGregor, K. K. (2017, June). *Closing the word-learning gap: Extra encoding opportunities equate long-term outcomes for students with and without developmental language disorders*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.
- McGregor, K. K., Arbisi-Kelm, T., & Eden, N. (2017). The encoding of word forms into memory may be challenging for college students with developmental language impairment. *International Journal of Speech-Language Pathology*, 19(1), 43–57.
- McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science*, 19(1), 22–24.
- 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.
- 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.
- Montgomery, J. W. (2000). Relation of working memory to off-line and real-time sentence processing in children with specific language impairment. *Applied Psycholinguistics*, 21(1), 117–148.
- 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.
- Montgomery, J. W., & Evans, J. L. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52(2), 269–288.
- Morice, R., & McNicol, D. (1985). The comprehension and production of complex syntax in schizophrenia. *Cortex*, 21(4), 567–580.
- Nation, K., Marshall, C. M., & Altmann, G. T. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology*, 86(4), 314–329.
- Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology*, 90(2), 142–171.
- Papesh, M. H., & Goldinger, S. D. (2012). Memory in motion: Movement dynamics reveal memory strength. *Psychonomic Bulletin & Review*, 19(5), 906–913.
- 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.
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Ryskin, R. A., Qi, Z., Duff, M. C., & Brown-Schmidt, S. (2017). Verb biases are shaped through lifelong learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(5), 781–794.
- Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49(3), 238–299.
- Spivey, M. J. (2007). *The continuity of mind*. New York, NY: Oxford University Press.
- Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Science of the United States of America*, 102, 10393–10398.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634.
- Tipper, S. P., Howard, L. A., & Houghton, G. (2000). Behavioral consequences of selection from population codes. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 223–245). Cambridge, MA: MIT Press.

Tomblin, J. B., & Nippold, M. A. (2014). Features of language impairment in the school years. In J. B. Tomblin & M. A. Nippold (Eds.), *Understanding individual differences in language development across the school years* (pp. 79–116). New York, NY: Psychology Press.

Trueswell, J. C., & Tanenhaus, M. (1994). Toward a lexical framework of constraint-based syntactic ambiguity resolution. In

C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–180). Hillsdale, NJ: Erlbaum.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(3), 528–553.

Appendix (p. 1 of 2)

Auditory (Numbered) and Visual (a: Instrument, b: Modifier, c: Incorrect) Stimuli

1. The bear chooses the duck with the fork.
 - a. bear using a fork to point toward a duck
 - b. bear pointing toward a duck that is holding a fork
 - c. bear holding a fork while pointing to a duck with his paw
2. The frog sings to the hamster with the funnel.
 - a. frog singing through a funnel to a hamster
 - b. frog singing to a hamster that is holding a funnel
 - c. frog holding a funnel while singing to a hamster
3. The dog looks at the monkey with the glass.
 - a. dog looking through a glass cup at a monkey
 - b. dog looking at a monkey that is holding a glass cup
 - c. dog holding a glass while looking at a monkey
4. The sheep listens to the cow with the tube.
 - a. sheep holding a tube to its ear next to a cow that is talking
 - b. sheep with one ear perked up sitting next to a cow that is holding a tube
 - c. sheep holding a tube next to a cow that is talking
5. The raccoon yells at the chipmunk with the funnel.
 - a. raccoon yelling through a funnel at a chipmunk
 - b. raccoon yelling at a chipmunk that is holding a funnel
 - c. raccoon holding a funnel while yelling at a chipmunk
6. The alligator finds the crab with the stick.
 - a. alligator holding a stick to lift a leaf covering a crab
 - b. alligator lifting a leaf to find a crab that is holding a stick
 - c. alligator holding a stick while lifting a leaf with claw to find a crab
7. The lion talks to the rabbit with the tube.
 - a. lion talking through a tube to a rabbit
 - b. lion talking to a rabbit that is holding a tube
 - c. lion holding a tube while talking to a rabbit
8. The gorilla hugs the cat with the blanket.
 - a. gorilla using a blanket to hug a cat
 - b. gorilla hugging a cat that is holding a blanket
 - c. gorilla holding a blanket with one arm while hugging a cat with the other
9. The giraffe brushes the zebra with the sponge.
 - a. giraffe using a sponge to brush a zebra's back
 - b. giraffe brushing a zebra that is holding a sponge in its mouth
 - c. giraffe holding a sponge with one hand while brushing a zebra's back with the other

Appendix (p. 2 of 2)

Auditory (Numbered) and Visual (a: Instrument, b: Modifier, c: Incorrect) Stimuli

10. The kangaroo cleans the panda with the t-shirt.

- a. kangaroo using a t-shirt to wipe dirt off a panda
- b. kangaroo wiping dirt from a panda wearing a t-shirt
- c. kangaroo holding a t-shirt with one paw while wiping dirt off a panda with the other

11. The butterfly hits the grasshopper with the flower.

- a. butterfly using a flower to (gently) hit a grasshopper
- b. butterfly using its wings to (gently) hit a grasshopper that is holding a flower
- c. butterfly holding a flower with one wing while (gently) hitting a grasshopper with the other

12. The squirrel tickles the skunk with the fan.

- a. squirrel using a fan to tickle a skunk
- b. squirrel tickling a skunk who is holding a fan
- c. squirrel holding a fan in one paw while tickling a skunk with the other

13. The elephant pokes the camel with the feather.

- a. elephant using a feather to poke a camel
- b. elephant using its trunk to poke a camel that is holding a feather
- c. elephant holding a feather with its mouth while poking a camel with its trunk

14. The tiger bops the hippopotamus with the ball.

- a. tiger throwing a ball at a hippopotamus
- b. tiger (gently) hitting a hippopotamus that is holding a ball
- c. tiger holding a ball with one paw while (gently) hitting a hippopotamus with the other

15. The owl covers the snail with the book.

- a. owl using a book to cover a snail
- b. owl using its wing to cover a snail that is reading a book
- c. owl holding a book in one wing while using the other to cover a snail

16. The deer feeds the fox with the glass.

- a. deer holding a glass cup to a fox's mouth
 - b. deer holding food to a fox's mouth while the fox holds a glass cup
 - c. deer holding a glass cup with one hoof while feeding a fox with another hoof
-