JSLHR

Research Article

Sentence Repetition Accuracy in Adults With Developmental Language Impairment: Interactions of Participant Capacities and Sentence Structures

Gerard H. Poll,[a] Carol A. Miller,[b] and Janet G. van Hell[b,c]

Purpose: We asked whether sentence repetition accuracy could be explained by interactions of participant processing limitations with the structures of the sentences. We also tested a prediction of the procedural deficit hypothesis (Ullman & Pierpont, 2005) that adjuncts are more difficult than arguments for individuals with developmental language impairment (DLI).
Method: Forty-four young adults participated, 21 with DLI. The sentence repetition task varied sentence length and the use of arguments and adjuncts. We also administered measures of working memory and processing speed. Our regression models focused on these interactions: group and argument status; processing speed, length, and argument status; and working memory capacity, length, and argument status.
Results: Language ability group was a significant predictor of sentence repetition accuracy but did not interact with argument status. Processing speed interacted with sentence length and argument status. Working memory capacity and its separate interactions with argument status and sentence length predicted sentence repetition accuracy.
Conclusions: Many adults with DLI may have difficulty with adjuncts as a result of their working memory limitations rather than their language ability. Cognitive limitations common to individuals with DLI are revealed more by particular sentence structures, suggesting ways to construct more diagnostically accurate sentence repetition tasks.

Sentence repetition is a widely used task for the assessment of developmental language impairment (DLI) as part of standardized tests (Hammill & Newcomer, 1997; Semel, Wiig, & Secord, 2003). Researchers have confirmed the diagnostic validity of sentence repetition tasks across age ranges (Conti-Ramsden, Botting, & Faragher, 2001; Poll, Betz, & Miller, 2010), languages (Devescovi & Caselli, 2007; Stokes, Wong, Fletcher, & Leonard, 2006; Vang Christensen & Hansson, 2012), and disability categories (Redmond, 2005; Taylor, Maybery, Grayndler, & Whitehouse, 2014). Because sentence repetition accuracy separates individuals with DLI from those with typical language, it has been designated a clinical marker of language impairment (Conti-Ramsden et al., 2001).

Clinical marker tasks for DLI, including tense marking, nonword repetition, and sentence repetition (Conti-Ramsden et al., 2001), have been advanced as both diagnostic tools and indicators of fundamental difficulties that result in language impairment. Rice and Wexler (1996) suggested that performance on tense marking tasks resulted from a deficit in grammatical knowledge, but others have attributed performance to processing limitations (Joanisse & Seidenberg, 1998; Leonard, 2014). Gathercole and Baddeley (1990) proposed nonword repetition as a measure of phonological memory, but contrasting evidence suggests that it measures phonological processing ability (Jones, Tamburelli, Watson, Gobet, & Pine, 2010). Sentence repetition has been advanced as a measure of the episodic buffer component of working memory (WM; Alloway, Gathercole, Willis, & Adams, 2004; Baddeley, Hitch, & Allen, 2009), but as with other clinic markers, there is evidence that other factors, such as difficulties processing complex syntax, contribute to sentence repetition accuracy (Riches, 2012). The purpose of

[a]Elmhurst College, Elmhurst, IL
[b]The Pennsylvania State University, University Park
[c]Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands

Correspondence to Gerard H. Poll: gerard.poll@elmhurst.edu

Editor: Rhea Paul
Associate Editor: Marleen Westerveld

Received January 22, 2015
Revision received May 21, 2015
Accepted August 12, 2015
DOI: 10.1044/2015_JSLHR-L-15-0020

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

this study was to explore how WM and other factors contribute to sentence repetition performance.

Sentence repetition performance may result from interactions of participant characteristics such as WM, processing speed, and language impairment status with sentence structure (Archibald & Joanisse, 2009; Hesketh & Conti-Ramsden, 2013). These interactions can suggest ways to make sentence repetition more diagnostically valuable. A recent meta-analysis found that existing sentence repetition tasks have only moderate diagnostic accuracy, finding a blended positive likelihood ratio of 6.87 across tasks (Pawlowska, 2014), with none meeting the standard of 10 or greater recommended by Dollaghan (2007). When a language-impaired group is disproportionately inaccurate in repeating sentences with a particular structure, tasks employing that structure will more clearly differentiate affected from unaffected individuals, improving diagnostic accuracy. When a characteristic of individuals with DLI, such as slow processing, interacts with sentence structure (Poll et al., 2013), it suggests that sentences may also be designed to reveal information about the profile of DLI when other correlated characteristics of individuals with DLI are well controlled.

Sentence Repetition

Sentence repetition is reconstructive (Potter & Lombardi, 1998; Slobin & Welch, 1971). Speakers generate sentences using their language production system, and do not simply play back what they just heard from auditory memory (Bley-Vroman & Chaudron, 1994; Lust, Flynn, & Foley, 1996). The sentences children produce reflect their level of language development (Devescovi & Caselli, 2007; Klem et al., 2015), and adults produce sentences that are influenced by primed words and structures (Lombardi & Potter, 1992; Potter & Lombardi, 1998).

The first step in sentence repetition is for the participant to hear and understand the target sentence. They then retain the essential meaning of the sentence in short-term memory, and engage their speech production system to produce a sentence on the basis of the concepts they are holding in mind (Bley-Vroman & Chaudron, 1994; Bock & Levelt, 1994; Potter & Lombardi, 1998). Participants recall more words presented in the form of a sentence than they can recall from a list of unrelated words (Baddeley et al., 2009; Jefferies, Ralph, & Baddeley, 2004). The reason for the superior recall of words in sentences is binding—more words can be bound together in larger units or chunks. This binding process results from the interaction of short-term stores of phonological information and language knowledge held in long-term memory. The episodic buffer in the Baddeley (2000) conception of WM is the limited capacity store for the chunks that result from the binding process. Clearly participants call on multiple elements of WM and long-term language knowledge in sentence repetition (Riches, 2012).

Sentence repetition has revealed gaps in knowledge of tense marking (Hesketh & Conti-Ramsden, 2013; Vang Christensen & Hansson, 2012) and difficulties with complex sentence structures in individuals with DLI (Panagos & Prelock, 1982; Redmond, Thompson, & Goldstein, 2011; Riches, Loucase, Baird, Charman, & Simonoff, 2010). Children with DLI have particular difficulties repeating sentences with complex structures such as object relatives (Riches et al., 2010). Sentence repetition accuracy is also compromised by WM limitations (Alloway & Gathercole, 2005; Archibald & Joanisse, 2009) and limitations in speed of language processing (Bishop, 1994; Poll et al., 2013).

Evidence for a role of individual differences in language processing speed in sentence repetition accuracy is limited, and suggests that processing speed limitations may not uniformly affect sentence repetition performance (Poll et al., 2013). The generalized slowing hypothesis suggests that processing speed in DLI may be slower by a fixed amount of time per cognitive operation (Kail, 1994). Sentences that require few cognitive operations may be affected very little by slower processing, but those requiring more cognitive operations may be more significantly affected.

Multiple conceptions of WM can be considered to understand its role in sentence repetition. In the Baddeley (2000) model, sentence recall is thought to measure the episodic buffer. In the Just and Carpenter (1992) model, WM is conceived as a storage and processing capacity that draws on a common mental resource. Sentence repetition involves comprehension, language processing, and storage of the meaning derived from the sentence. This capacity, measured by the Competing Language Processing Task (CLPT; Gaulin & Campbell, 1994), improves during development and is limited in individuals with DLI (Ellis Weismer, Evans, & Hesketh, 1999; Isaki, Spaulding, & Plante, 2008).

Both the Baddeley (2000) and Just and Carpenter (1992) conceptions of WM have been measured using complex tasks that involve trade-offs between storage and processing. The Just and Carpenter (1992) model focuses on sentence comprehension and the dynamic trade-off between storage and processing as sentences vary in complexity, making it an apt focus for understanding sentence repetition. A third conception of WM is how much information an individual can maintain in the focus of their attention, emphasizing storage capacity (Cowan et al., 2005). Measures of this notion of WM have not been reported for adults with DLI, but such data would clarify the importance of the storage element of WM in the profile of the disorder.

Processing and Linguistic Knowledge Differences in DLI

The procedural deficit hypothesis (PDH; Ullman & Pierpont, 2005) is a theory of DLI that integrates limitations in language knowledge and limitations in processing capacity, both of which play roles in sentence repetition accuracy. The PDH builds on a theory of language processing that separates grammatical and lexical processing (Ullman, 2001). By this theory, grammatical processing relies on rules, whereas lexical processing relies on associative memory.

Procedural memory supports implicit learning of sequences, skills, and patterns over time (Squire, 1992; Tulving, 1985). For language, procedural memory underlies syntactic learning and processing. Declarative memory supports explicit learning and recall of facts and events (Tulving, 1985), as well as the idiosyncratic associations of words, their referents, and obligatory arguments (Ullman, 2001). The PDH proposes that a deficit in procedural memory results in difficulties with syntactic processing (Ullman & Pierpont, 2005). The declarative memory system of individuals with DLI is thought to be intact or a relative strength.

Studies have shown that children (Kemeny & Lukacs, 2010; Lum, Gelgic, & Conti-Ramsden, 2010), adolescents (Tomblin, Mainela-Arnold, & Zhang, 2007), and adults (Lee & Tomblin, 2012) with DLI have more limited procedural memory ability for verbal information. Findings are less consistent with the PDH when the procedural memory task involves nonverbal information (Lum & Conti-Ramsden, 2013) or when the task does not involve learning discrete sequences (Hsu & Bishop, 2014). Given proximity of brain circuits supporting procedural memory with those supporting WM, the PDH anticipates that individuals with DLI will have difficulty with tasks demanding of WM (Ullman & Pierpont, 2005). Evidence on declarative memory for individuals with DLI is mixed, finding intact memory for visual material (Lum & Conti-Ramsden, 2013; Riccio, Cash, & Cohen, 2007) and deficits for verbal material (Lum, Conti-Ramsden, & Ullman, 2012; Riccio et al., 2007).

An untested prediction of the PDH is that individuals with DLI will have unusual difficulties with adjuncts and less difficulty with arguments. Arguments and adjuncts, as sentence constituents, differ in their reliance on lexical processing versus syntactic processing. Across multiple theories of syntax, arguments are viewed as part of lexical knowledge (Chomsky, 1970; Quirk, Greenbaum, Leech, & Svartvik, 1985; Radford, 2004). In *The president urged the army to act*, the verb *urge* implies an agent to carry out the action (*the president*) and a recipient of the action (*the army*). The agent and recipient act as arguments of the verb. Because arguments are implied by the meaning of the word, arguments are part of a word's lexical entry in the mental dictionary. That lexical information eases the integration of arguments into sentence structures during comprehension (van Gompel & Pickering, 2007).

Adjuncts, or modifiers, are peripheral to the meaning of the words they modify (Quirk et al., 1985; Radford, 2004; Thompson, 1997). In *The queen visited before the judgment*, the phrase *before the judgment* is an adjunct, an optional modifier specifying when the visit took place. Understanding the meaning of an adjunct phrase in a sentence requires that the phrase be combined with the phrase it modifies (here, *visited*), a syntactic operation (Boland & Boehm-Jernigan, 1998) that is made more difficult by the absence of adjunct information in the verb's lexical entry (Boland & Blodgett, 2006).

The PDH predicts that individuals with DLI will have unusual difficulties processing adjuncts because the syntactic processing they require depends on a deficient procedural memory (Ullman & Pierpont, 2005). Studies of children with DLI have shown that they are more likely to omit adjuncts that are included by typical peers (Fletcher & Garman, 1988; Johnston & Kamhi, 1984). The PDH also suggests that declarative memory may compensate for procedural deficits (Ullman & Pierpont, 2005) but that the ability to compensate will be sensitive to the frequency of language structures. Because adjuncts tend to co-occur less frequently with the words they modify than do arguments (Boland & Blodgett, 2006), it is unclear how well declarative memory can compensate for procedural deficits in processing adjuncts.

In contrast, the PDH predicts that individuals with DLI should be able to process arguments much as individuals with typical language do because argument knowledge relies on their intact declarative memory (Ullman & Pierpont, 2005). For children with DLI, some studies have supported typical or near-typical processing of arguments (Grela & Leonard, 2000; Thordardottir & Ellis Weismer, 2002), although there is counterevidence suggesting more complex argument structures are particularly difficult (de Jong & Fletcher, 2014). It is unclear whether these patterns of language processing apply to adults with DLI.

DLI in Adulthood

Multiple studies have documented the chronic nature of DLI, and demonstrated that DLI extends into adulthood (Clegg, Hollis, Mawhood, & Rutter, 2005; Johnson et al., 1999). Understanding the adult stage of DLI is clinically important to address the transition of adolescents with DLI to postsecondary education (Sitlington, 2003) and to support adults with DLI in postsecondary schools and vocational settings.

Sentence repetition provides a promising assessment tool for identifying adults with DLI (Poll et al., 2010; Tomblin, Freese, & Records, 1992) for the same reasons that it is useful for children with DLI. It taxes WM, processing speed, and syntactic knowledge, all of which appear to be compromised in adults with DLI (Isaki et al., 2008; Miller & Poll, 2009; Poll et al., 2010; Tomblin et al., 1992).

During development, language and cognitive abilities interact, resulting in relatively lower nonverbal intelligence for those with DLI as compared with peers with typical language (Botting, 2005). As a result, studies have not matched adults with DLI to peers with typical language on nonverbal intelligence (Fidler, Plante, & Vance, 2011; Lee & Tomblin, 2012; Tomblin et al., 1992). In fact, researchers have argued that matching or statistically controlling for nonverbal ability may result in groups that are no longer representative of the populations they are meant to represent (Dennis et al., 2009).

Given that adults with DLI have limitations in both processing capacities and syntactic abilities, we sought to understand how both participant abilities and sentence structure might explain sentence repetition accuracy. We tested the prediction of the PDH that adjuncts will pose particular difficulties for adults with DLI. Our sentences varied by being primarily composed of arguments or adjuncts.

A significant Group × Argument Status interaction, with less accurate adjunct sentence scores in the group with DLI, would support the prediction of the PDH. We then evaluated how processing limitations interact with sentence structures in sentence repetition across all participants. Is the effect of slower language processing greater for sentences requiring more cognitive operations? Does the effect of more limited WM capacity vary according to the structure of the sentence?

Method

Participants

Forty-four adults (age 18–27 years) participated in the study and were recruited from postsecondary schools in central Pennsylvania (n = 22) and from a registry of participants in an Iowa longitudinal study of DLI (n = 22). Initial selection of the Iowa participants is described in Tomblin, Zhang, Buckwalter, and O'Brien (2003). We excluded participants if they reported a history of autism, intellectual disability, cerebral palsy, hearing loss, or significant neurological injury. All participants spoke English as a first language.

We screened participants for hearing loss with a puretone screening at 25 dB HL at 500, 1000, 2000, and 4000 Hz. Participants completed three subtests of the Wechsler Adult Intelligence Scale–Third Edition (Wechsler, 1997), Picture Completion, Digit Symbol Coding and Matrix Reasoning. An estimated performance intelligence quotient (PIQ) was calculated from the scores of these subtests based on Sattler and Ryan (1999). All participants included in the study had PIQs of 80 or above.

Screened participants were assigned to a language group on the basis of a positive or negative history of language difficulties and language testing. Those in the group with DLI had a positive history (prior diagnosis of DLI, reading comprehension difficulties, or spoken grammar difficulties). Those in the group with typical language had no history of language difficulties. Each group included participants from both the Pennsylvania (DLI = 7, typical = 15) and Iowa (DLI = 14, typical = 8) recruitments.

Language testing followed the approach outlined by Fidler et al. (2011). Their diagnostic process has the best documented sensitivity (78%) and specificity (83%) of any assessment for DLI in English-speaking young adults. The tasks were the Modified Token Test (de Renzi & Faglioni, 1978; Morice & McNicol, 1985), the Word Definitions subtest of the Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003), and a 15-word spelling task (Fidler et al., 2011). The Modified Token Test consists of sentences directing participants to manipulate tokens of varied shape and color. The score is the number correct. The Word Definitions subtest asks participants to provide spoken word meanings. Scores were according to test directions, converted to standard scores for 18- to 21-year-olds. Eleven percent of Word Definitions subtests were scored by a trained second rater, with 88% point-to-point agreement. Scores for the spelling task were the number correct.

Scores from the three language tests were entered into a discriminant function developed by Fidler et al. (2011) to diagnose DLI in adults with a history of learning difficulties. A positive numeric output indicates that the person is language impaired. The 21 participants with a positive history of language difficulties and a positive discriminant function output were classified as DLI. The 23 participants with no history of language difficulties and a negative discriminant function output were classified as typical language. Participants with a positive history and a negative discriminant function output were excluded. Group characteristics are summarized in Table 1. There were 13 women in the group with DLI, and 19 women in the group with typical language. The groups did not differ in age, $F(1, 42) = 2.83$, $p = .10$, but did differ in years of education and PIQ.

Materials and Procedures

Sentence Repetition Task

We developed 48 sentences in four conditions of 12 sentences each. The sentences were short (16 syllables) or long (25 syllables), and were argument or adjunct laden. We controlled length by syllables rather than words because adults with DLI have more difficulty with memory for longer compared with shorter words (Clegg et al., 2005). We classified phrases as arguments or adjuncts using tests developed by Schutze and Gibson (1999). Their tests rely on native speaker judgments of grammaticality. Among the tests for argumenthood is optionality: adjuncts can be dropped and the sentence remains grammatical (e.g., *John put the book in the room* is grammatical; *John put the book* is not). Another test is iterativity: arguments cannot be iterated whereas adjuncts can (e.g., *John amused the earl with his description* is grammatical; *John amused the earl with his description with his music* is not). The complete set of experimental sentences is provided in the online supplemental materials (see Supplemental Text).

Sentences in argument conditions included verbs that are more often transitive, meaning a noun phrase argument is likely to follow the verb. A short argument condition sentence was *The dentist baked her assistant a cake and added a cherry*. Here *baked* is a verb that is often transitive. Adjunct condition sentences included intransitive verbs, those not likely to require a following object noun phrases. A short adjunct condition sentence was *The queen visited before the judgment but prayed at great length*. Here *visited* and *prayed* are verbs more likely to be intransitive. We identified transitive or intransitive verbs from corpus studies (Gahl, Jurafsky, & Roland, 2004; Roland, Dick, & Elman, 2007).

Sentences in short conditions had a mean of two functional verb units, or two finite or nonfinite verbs (Bock & Levelt, 1994; Ford & Holmes, 1978). Sentences in long adjunct conditions also typically had two functional verb units, such as *confided* and *surrendered* in *The couple confided a breakdown in the spring and surrendered their house in the valley after the split*. The mean number of functional verb units for the long adjunct condition was 2.17. Sentences in long argument conditions had a mean of three functional verb

Table 1. Participant group characteristics.

Measure	Group with DLI		Group with TL		d^a
	M (SD)	Minimum–maximum	*M (SD)*	Minimum–maximum	
n	21		23		
Age	22.4 (2.0)	18–27	21.5 (1.8)	18–25	.47
Years of education	13.1 (1.1)	11–14	14.5* (0.9)	14–16	1.39
Performance IQ	97.5 (7.9)	86–121	113.7* (10.0)	94–131	1.80
Modified Token Test	70.4 (17.1)	25–98	91.3* (5.1)	77–100	1.66
Spelling	3.7 (2.2)	0–7	11.4* (1.9)	9–15	3.75
Word Definitions	7.7 (3.3)	3–13	13.1* (1.5)	9–15	2.11

Note. The performance IQ is calculated from the Picture Completion, Digit Symbol Coding, and Matrix Reasoning subtests of the Wechsler Adult Intelligence Scale–Third Edition (Wechsler, 1997) using the approach from Sattler and Ryan (1999). The Modified Token Test (de Renzi & Faglioni, 1978; Morice & McNicol, 1985) scores are the group mean percentage correct. Spelling is the number of words spelled correctly of the 15 words presented from Fidler, Plante, and Vance (2011). Word Definitions is the group mean standard score for the subtest from the Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003). DLI, developmental language impairment; TL, typical language.

[a]Effect size of the group mean difference.

*$p < .05$ for group mean difference.

units. The number of functional verb units is a measure of the number of cognitive operations carried out in formulating a sentence (Ford & Holmes, 1978).

Sentences were controlled for word frequency, word familiarity, and sentence plausibility. Spoken word frequency for the subject nouns, the verbs, the object nouns, and the modifiers did not differ by condition, $F(3, 305) = 0.88$, $p = .452$. We obtained frequency from the Corpus of Contemporary American English (Davies, 2009). Word familiarity and sentence plausibility were rated by 14 adults with typical language (age 18–25 years), all speakers of English as a first language. Participants were asked how often they had seen, heard, or used a word using a 7-point scale. They were also asked to rate how likely sentence events were to occur using a 7-point scale. Neither word familiarity ratings, $F(3, 303) = .65$, $p = .584$, nor sentence plausibility ratings, $F(3, 44) = 1.51$, $p = .225$, differed across sentence repetition task conditions.

Sentences were pseudorandomly assigned to two sets, each with equal numbers of sentences from each of the four conditions. Sentences in each set were presented in three different lists in random orders. Each participant repeated all 48 sentences but did so in two different sessions separated by a minimum break of one-half hour. Sentence lists were counterbalanced across participants.

Sentences were digitally audio-recorded by a male speaker of American English and edited to a uniform intensity level using Praat (Boersma & Weenink, 2006). We presented sentence recordings by computer. We instructed participants to repeat the sentences exactly as they heard them, and although some sentences were long and would be hard to remember, they were to repeat as much of the sentence as they could. Participants completed two practice sentences before the experimental task. An investigator marked which words the participant correctly repeated at the time of the task, and we verified which words had been repeated correctly from recordings.

Scores were percentage accuracy, the number of words recalled in correct serial order, divided by the number of words in the sentence. Following Allen and Baddeley (2009), we credited correct initial or final words in sentences, as well as any word with at least one adjacent word in correct serial position. We preferred this method over the Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003) 0 to 4 rating system because it did not result in floor effects. A second rater with no knowledge of participant group classification was trained on the scoring procedure and scored 11% of the sentences. The intraclass correlation was .98, suggesting that very little variation in the scores was due to rater variation (Howell, 2007).

Processing Speed Tasks

To measure language processing speed, we administered a truth value judgment task and a word detection task. For the truth value judgment task, participants saw a line drawing presented on a computer, and after 2 s heard a sentence. They were asked to press one button on a response box to indicate that the picture and sentence matched and another to indicate that they did not. The task included 36 pictures and sentences from Miller, Kail, Leonard, and Tomblin (2001). The sentences consisted of simple active (*The girl is chasing the boy*), passive (*The baby is being fed by the girl*), or compound subject (*The horse and the dog are chasing the cat*) constructions. The sentences and pictures were evenly divided between matching and not matching conditions.

Sentences were audio-recorded by a male speaker of American English. The task was by computer using E-Prime (Psychology Software Tools, 2009). Picture and sentence pairs were presented in a single pseudorandom order, with no more than three consecutive matching or nonmatching pairs. Participants indicated their judgments using a response box aligned with their dominant hand. Participants wore headphones, and sentences were presented at a comfortable loudness. We instructed participants to respond as quickly as possible without making mistakes. There were two practice sentences, during which the investigator reinforced the instructions.

Scores for the truth value judgment task were mean response times in milliseconds across all trials. Response time was measured from the start of the sentence to the button press, so the total response time included time to listen to the sentence. Incorrect judgments and response times more than twice the participant's mean were excluded as outliers.

For the word detection task, participants heard a target word, then a sentence. Participants were to press a button as soon as they heard the target word in the sentence. There were 34 target words and related grammatical sentences drawn from a task described in Leonard, Miller, and Finneran (2009). Sentences were eight to 12 words, and the target words appeared as the fifth, sixth, or seventh words. In nine sentences, the target word did not appear. These sentences were included to encourage participants to listen carefully to the entire sentence.

Target words appeared at least twice in a written language corpus (Francis & Kucera, 1982). Target words and sentences were recorded by a male speaker of American English and adjusted to a common intensity. The task was presented by a computer using E-Prime (Psychology Software Tools, 2009) in two pseudorandom orders, counterbalanced across participants. Sentences not including the target word occurred at least once in each nine sentences. Sentences analyzed for this study were presented at a normal speaking rate. An equal number of sentences presented at a slow rate were included in the task, but those results were not a part of the study reported here. Instructions for the word monitoring task were to press a button as soon as the participant heard the target word in the sentence, or if they did not hear the target word, to do nothing. The task was presented at a comfortable loudness using headphones. Each trial started with 2 s of silence followed by the target word. After 500 ms, the sentence sound file was played. Trials ended after 2 s or a participant response.

The score for the word detection task was the mean response time from the onset of the target word in the sentence to the participant button press. Negative response times, indicating that the participant pressed the button before the target word occurred, were excluded. We also excluded response times greater than twice the participant's mean as outliers.

WM Tasks

We evaluated WM using the CLPT (Gaulin & Campbell, 1994) as a storage and processing measure and Running Span (Cowan et al., 2005) as a storage-focused measure. The CLPT asks participants to listen to sets of sentences and render a truth judgment for each. After hearing a set of sentences, they are asked to recall the last word of each sentence in the set. There are 42 sentences arranged in sets of one to six sentences. All sentences are simple sentences such as *Pumpkins are purple*. We presented the task as described in Gaulin and Campbell (1994). Prior to the task, we gave instructions from Ellis Weismer et al. (1999). In the task recording, instructions were repeated according to the script in Gaulin and Campbell (1994).

The task was presented at a comfortable loudness under headphones using a recording by a female speaker of American English. The recording allowed 4 s for truth judgments and 7 s (one-sentence sets) to 19 s (six-sentence sets) to recall the final words. Participants were allowed only this pause time to provide responses, following Mainela-Arnold and Evans (2005). The task began with four practice sentences. We noted truth–value judgments and final words recalled at the time of administration, and verified the accuracy from audio recordings. The score for the task was the percentage of sentence-final words correctly recalled. Morphological variations were accepted as correct.

Running Span

In the Running Span task, participants heard lists of digits. The lists varied in length from 12 to 20 digits. After hearing each list, the participant was asked to recall the last five, six, or seven digits in the list in the correct forward order. This task was modeled on the Running Span task described in Cowan et al. (2005).

The numbers one to nine were recorded one at a time by a male speaker of American English. The sound files were edited and compressed to 250 ms with Praat (Boersma & Weenink, 2006). The task was presented using a computer running E-Prime (Psychology Software Tools, 2009). The task began with a practice block of two lists of digits. After hearing each list, the participant was asked to write the last five digits from the end of the list onto a response form. The response form had five boxes for the response. The participant was instructed to fill all of the boxes, and use zero if they could not recall a digit, because zero was not included in the lists. The investigator provided feedback and repeated instructions to participants throughout the practice block.

There were three experimental blocks, each consisting of nine lists of digits, with each list length randomly varied from 12 to 20 digits. The computer script randomly selected digital audio files without replacement, and then repeated the process in order to complete each list. For the first experimental block, the participant wrote the final seven digits from the list. For the second block, they wrote the last six, and for the final block, the last five digits.

Each trial began with the word *Ready* on the screen. After pressing the space bar, participants heard the digit list followed by presentation of response boxes. They wrote their responses on paper, and the investigator transferred these to the computer. There was no time limit for responses. Participants verified that what was entered into the computer matched their paper response. Pilot participants indicated that writing the response on paper rather than keying it into a computer made the task easier.

The computer script created a record of the final seven, six, or five digits presented and the participant response. The score for each trial was the number of digits recalled in correct order. The mean number correct was calculated for each block, and the highest block mean was the participant score.

Analysis Approach

In order to identify predictors of sentence repetition accuracy, we developed regression models with the mean sentence repetition scores of each participant in each condition as the dependent variables. Because item-level scores by participant in the typical group had ceiling effects, we used by-participant condition means. These had distributions that met the assumptions for regression modeling.

To account for the repeated measures (within participant) element of our design, we used mixed-effects regression models. These models included random effects terms for participants (Baayen, 2008). Inclusion of random effects for intercepts only or intercepts and slopes can be evaluated to determine whether they improve the model fit to the data. Predictors of sentence repetition accuracy were entered as fixed effects. In the absence of consensus on how to compute degrees of freedom for these predictors, we determined statistical significance by whether the *t* statistic for the predictor was greater than 2.00, and by use of likelihood ratio tests (Arnon & Snider, 2010; Baayen, 2008). Likelihood ratio tests evaluate whether a model including a predictor has a significantly better fit to the data than a model excluding the predictor, using a chi-squared statistic (Tabachnick & Fidell, 2007).

Results

Sentence repetition mean accuracy by group and by condition is presented in Table 2. There were group differences in accuracy for all conditions with larger effect sizes for long sentences than for short, and for adjunct sentences compared with argument sentences.

Interactions of Group and Argument Status

We developed a mixed-effects regression model to evaluate a prediction of the PDH. We first entered a random by-participant intercept, and added a random slope by-participant by-length. The random slope improved model fit, $\chi^2(2) = 131.8$, $p < .001$, and so it was retained in the model.

The PDH predicts that individuals with DLI will have particular difficulties with adjunct processing. We entered fixed effects of group, argument status, and length together with their interactions into the model. Neither the Group × Argument Status, $t = 1.3$, $\chi^2(1) = 1.6$, $p = .204$, nor the Group × Length interactions, $t = 1.7$, $\chi^2(1) = 3.04$, $p = .081$, were significant. The main effects of group, $t = 5.5$, $\chi^2(3) = 26.5$, $p < .001$, argument status, $t = 3.3$, $\chi^2(3) = 47.8$, $p < .001$, and length, $t = 22.6$, $\chi^2(3) = 153.4$, $p < .001$, were significant predictors of sentence repetition accuracy. The Argument × Length interaction was also significant, $t = 5.5$, $\chi^2(3) = 26.5$, $p < .001$. The addition of years of education as a control variable to the model did not change the findings: The group, argument, and length predictors together with the argument by length interaction remained significant (all *t* values greater than 2.0).

The hypothesis from the PDH was that the group with DLI would have particular difficulty with adjunct conditions, resulting in a Group × Argument Status interaction. This hypothesis was not supported. The model confirms that the group with DLI was less accurate than the group with typical language. Both argument status and length explained sentence repetition accuracy, but the Argument × Length interaction indicated that the effect of argument status differed in short and long conditions. Participants were more accurate repeating short argument-laden sentences as compared with short adjunct-laden sentences. In long conditions, the argument sentences did not have an accuracy advantage.

Processing Capacity and Sentence Structure Interactions

We next evaluated how processing capacity affected sentence repetition accuracy. Participant scores for the WM

Table 2. Group performance on sentence repetition, working memory, and processing speed tasks.

	DLI		TL		
Measure	***M (SD)***	**Minimum–maximum**	***M (SD)***	**Minimum–maximum**	***d*[a]**
Sentence repetition by condition					
Short argument sentences	82.2 (21.2)	52.2–96.3	94.9 (11.6)	78.8–100.0	0.74*
Short adjunct sentences	69.9 (24.8)	41.3–87.4	87.7 (19.3)	69.4–99.3	0.81*
Long argument sentences	38.1 (20.3)	24.5–66.7	58.7 (26.1)	33.4–98.6	0.88*
Long adjunct sentences	40.0 (22.7)	20.8–78.7	63.3 (24.9)	34.2–97.9	0.98*
Working memory measures					
CLPT recall percentage	69.9 (11)	40–88	86.4 (10)	69–100	1.6*
Running Span digits	3.2 (.62)	2.1–4.6	4.2 (.73)	3.2–6.3	1.5*
Processing speed measures					
Truth Value RT (ms)	2,358 (292)	2,019–3,147	2,199 (304)	1,662–2,917	0.5
Word Detection RT (ms)	431 (126)	275–785	341 (60)	257–488	0.9*

Note. Sentence repetition scores are percentage of words correctly repeated. DLI, developmental language impairment; TL, typical language; CLPT, Competing Language Processing Task (Gaulin & Campbell, 1994); Running Span digits, the maximum number of digits recalled for any block; Truth Value RT, response time for indicating whether a sentence and picture matched; Word Detection RT, response time for indicating that a target word appeared in a sentence.

[a]Effect size of the group mean difference.

*Test of group mean differences $p < .05$.

and processing speed tasks are presented in Table 2. With the exception of the truth value judgment task, there were significant group differences in task performance, with large effect sizes. Table 3 summarizes for each group the correlations between processing capacity measures and sentence repetition mean accuracy across conditions. Processing speed measures did not correlate with overall sentence repetition accuracy, whereas for the group with typical language, WM measures did.

Our analysis of the effects of processing capacity took an individual differences approach—we considered the effect of processing capacity across the full range of our participants. We first evaluated the role of processing speed. We calculated a composite processing speed score by summing the *z* scores of the word detection and truth value judgment tasks. We entered the processing speed composite into a mixed-effects model to predict sentence repetition accuracy. The random intercepts by participant and random slope by length were included as in the previous model. The composite processing speed variable was not a reliable predictor, $t = 1.96$, $\chi^2(1) = 3.82$, $p = .051$. We then entered the standardized word detection response time into the regression model in place of the processing speed composite. Word detection response time predicted sentence repetition accuracy, $t = 2.56$, $\chi^2(1) = 6.38$, $p = .01$.

We added the argument status and length factors and their interactions with word detection response time to the model, including a three-way interaction of argument status, length, and word detection response time. The three-way interaction allowed us to evaluate whether processing speed was a larger factor in sentence repetition accuracy when sentences required more cognitive operations, operationalized as functional verb units. The long argument condition sentences had more functional verb units than the other conditions. Length, argument status, and word detection response time were significant predictors (*t*s > 2.00). Interactions of length by argument status ($t = 7.03$) and length by argument status by word detection response time, $t = 2.15$, $\chi^2(1) = 4.72$, $p = .03$, were significant. The two-way interactions of word detection response time with length ($t = .27$) and argument status ($t = .41$) were not significant. The three-way interaction remained significant with the addition of years of education as a control variable ($t = 2.15$).

The significant three-way interaction of length, argument, and processing speed indicates that slower processing affected sentence repetition differently in one sentence condition. We divided the participants into above and below median word detection response time groups and found that the accuracy advantage of the faster processors was six percentage points for all conditions except the long argument condition, in which the advantage was 7.5 percentage points. Processing speed had a larger effect on the condition made up of sentences with the largest number of functional verb units.

We next evaluated whether the effect of WM capacity varied by sentence condition. Again we took an individual differences approach and did not form groups for the analysis. We entered a composite WM score (the sum of the standardized scores from the CLPT and Running Span) in a mixed-effects regression model predicting sentence repetition accuracy. We included the same random effects structure as in the models described above. The WM composite reliably predicted sentence repetition accuracy, $t = 6.56$, $\chi^2(1) = 24.0$, $p < .001$.

Adding length, argument status, and interactions to the model, we found that the two-way interactions of WM with length, $t = 4.17$, $\chi^2(1) = 15.24$, $p < .001$, and WM with argument status, $t = 2.33$, $\chi^2(1) = 5.38$, $p = .02$, were significant. The three-way interaction of WM by length by argument status was not, $t = 1.65$, $\chi^2(1) = 2.82$, $p = .09$. Main effects of argument status, length, WM composite, and the Argument × Length interaction were all significant predictors as well (*t*s > 2.00). In a model with years of education added as a control variable, the WM composite and the two-way interactions of WM with argument status and with length remained significant.

Figure 1 illustrates the two-way interactions of WM and argument status and of WM and length. Sentence repetition accuracy was more sensitive to WM capacity in long and adjunct conditions.

Table 3. Correlations of language, performance intelligence quotient (PIQ), processing capacity measures, and mean sentence repetition accuracy collapsed across conditions. Results for the group with developmental language impairment are above the diagonal. Results for the group with typical language are below the diagonal.

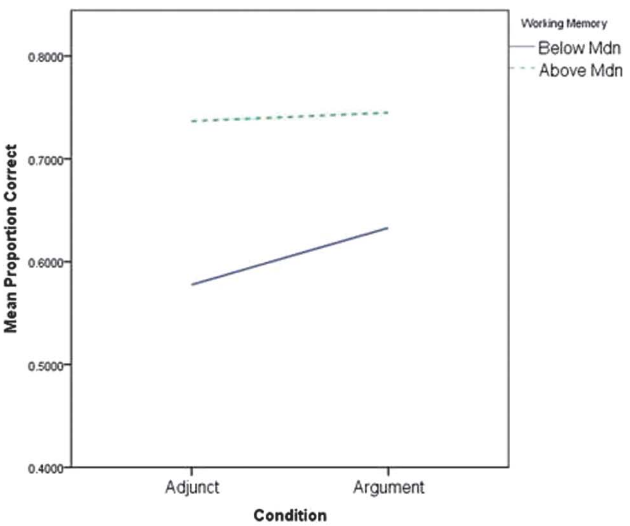
Measure	1	2	3	4	5	6	7
1. PIQ	1	−.19	.26	.33	−.12	−.45*	.28
2. Language[a]	−.51*	1	.12	−.25	−.15	−.24	−.44*
3. CLPT Recall	.36	−.58*	1	.34	.22	−.16	.25
4. Running Span	−.11	−.39	.51*	1	.07	−.19	.24
5. Truth Value RT	−.41	−.02	.18	.09	1	.21	.06
6. Word Detection RT	−.54*	.16	−.02	−.14	.47*	1	−.09
7. Sentence Repetition	.29	−.66*	.68*	.66*	−.03	−.08	1

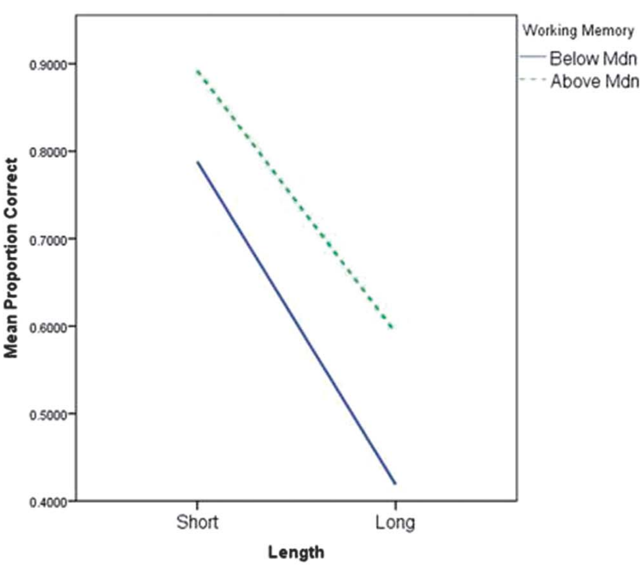
Note. CLPT, Competing Language Processing Task (Gaulin & Campbell, 1994); RT, response time.

[a]Output of the discriminant function specified in Fidler, Plante, and Vance (2011).

**p* < .05 for two-tailed test.

Figure 1. Participants split into working memory capacity above and below study median, with sentence repetition accuracy by argument versus adjunct condition (top) and short versus long conditions (bottom).





Group Differences in PIQ

Unlike our analyses of processing speed and WM, our model evaluating the PDH focused on the Group × Argument Status interaction, and so depended on group classification. The group with DLI differed from the group with typical language in both language ability and PIQ. This has frequently been the case in studies of adults with DLI (Fidler et al., 2011; Lee & Tomblin, 2012; Rost & McGregor, 2012; Tomblin et al., 1992). Because somewhat lower PIQ is characteristic of adults with DLI, matching on PIQ risks creating groups that are not representative of the population. Statistically controlling for PIQ results in similar issues (Dennis et al., 2009). It is important to know, however, whether our findings reflect primarily the PIQ difference or the language ability difference.

The correlations in Table 3 suggest that the language ability difference, rather than the PIQ difference, was the primary reason for our finding. We used the discriminant function output from our group classification process as a proxy for language ability. The correlations between language ability and sentence imitation were significant for both groups and stronger than the nonsignificant correlations between PIQ and sentence imitation accuracy. PIQ and language ability were, however, correlated for the group with typical language. We performed a sequential linear regression predicting the cross-conditional mean sentence imitation accuracy. We entered PIQ as the first predictor followed by language ability (discriminant function output). After entering language ability, PIQ was no longer a significant predictor of sentence imitation accuracy ($t = .72$, $p = .47$), but language ability was ($t = 4.71$, $p < .001$). This provides additional evidence that our finding of a significant group effect in our analysis evaluating the PDH prediction was primarily a result of the language ability difference between groups.

Discussion

Our goal in this study was to develop a better understanding of sentence repetition as a tool for the assessment of DLI. Although prior studies have identified factors involved in sentence repetition performance, our design allowed us to observe the interactions of processing limitations and sentence structures. We also tested a prediction of the PDH (Ullman & Pierpont, 2005).

The PDH

The PDH predicts difficulties with adjunct processing for individuals with DLI (Ullman & Pierpont, 2005). We contrasted adjunct-laden with argument-laden sentences, and found no significant Group × Argument Status interaction. All participants found adjuncts more difficult in short conditions, and individuals with DLI were not differentially affected by adjuncts. The PDH suggests that individuals with DLI may compensate for their difficulties with syntax by relying on declarative memory. They may learn associations of verbs with their adjunct modifiers that typical peers process using an implicit rule of grammar (Boland & Boehm-Jernigan, 1998). Because we did not have a measure of indicating use of declarative memory for this task, we were unable to evaluate this possibility.

An alternate explanation for the lack of a Group × Argument Status interaction is that adjunct processing does not categorically differ from argument processing in its reliance on syntactic ability. The PDH assumes that arguments are part of lexical knowledge associated with verbs whereas adjuncts are not (Ullman & Pierpont, 2005). Evidence from the sentence processing literature has generally supported that arguments are easier to process than adjuncts, but it is less certain whether this is due to a categorical distinction between arguments and adjuncts (Boland & Blodgett, 2006; Schutze & Gibson, 1999). Instead, arguments and adjuncts may differ in a graded way, depending on frequency with which they co-occur with their related verbs (MacDonald, Pearlmutter, & Seidenberg, 1994). If the distinction is purely one of frequency, then the argument–adjunct distinction in our materials may not have clearly differentiated syntactic and lexical processing requirements in our task conditions.

Prior evidence for adjunct difficulties in DLI comes from studies of children (Fletcher & Garman, 1988; Schuele & Tolbert, 2001). It is possible that at adulthood, individuals with DLI no longer have a special vulnerability to adjunct processing. Or, it may be that the special difficulty with adjuncts is not a characteristic of the general language ability difference in DLI, but is instead a result of the processing limitations characteristic of many individuals with DLI. The interaction of WM and argument status found in our data supports this view.

Speed of Processing Interactions With Sentence Structure

Slower language processing is characteristic of children and adults with DLI (Miller et al., 2001, 2006; Miller & Poll, 2009), but the degree of slowing has not always been a correlate of language ability measures (Lahey, Edwards, & Munson, 2001). Prior work has considered the generalized slowing hypothesis (Kail, 1994) and how it might explain the relation of deficits in processing speed to language tasks (Poll et al., 2013). If processing is slowed by a given amount for each cognitive operation, then language tasks requiring more operations will be more affected by a deficit in processing speed than those requiring fewer.

We found a three-way interaction of processing speed (word detection response time), argument status, and length. Means by condition and by above and below median processing speeds revealed that slower processors had a larger performance decrement for long argument conditions as compared with the other three conditions. Long argument conditions had the highest mean number of functional verb units, a proxy for number of planning operations required for sentence production (Ford & Holmes, 1978). The effect of slower processing was magnified in the condition that required more cognitive operations, consistent with the generalized slowing hypothesis (Kail, 1994). The long argument condition also included more complex verb argument structures than the other conditions. Generating plausible argument-laden sentences required the use of verbs that take more arguments, such as ditransitive verbs that take two arguments in addition to the agent. There is evidence that children with DLI have more difficulty with ditransitives than transitive or intransitive structures (Llorenc, Sanz-Torrent, Guardia Olmos, & MacWhinney, 2013; Thordardottir & Ellis Weismer, 2002). Children also require more planning time when formulating sentences with ditransitives (Llorenc et al., 2013), suggesting that argument structure complexity may be a proxy for the number of cognitive operations involved in formulating a sentence. In any case, adults with slower processing had unusual difficulty with long argument condition sentences.

A question raised by our findings was why the truth value judgment task did not result in group differences or correlations with sentence repetition accuracy. The truth value judgment task involved both visual processing and sentence comprehension. It is possible that the task, originally designed for children, was not challenging enough to reveal processing speed differences in adults.

WM Interactions With Sentence Structure

Unlike the processing speed measures, both WM measures resulted in reliable group differences and were together significant predictors of sentence repetition accuracy. A significant group difference resulting from our storage and processing measure, the CLPT, confirms differences in adults with DLI found in prior work (Isaki et al., 2008). The similarly large group difference on Running Span adds to the profile of WM limitations in adults with DLI. The results suggest that storage capacity, specifically how many items can be maintained in the focus of attention, is more limited among adults with DLI.

Simple tasks measuring short-term memory (such as forward digit span) have not reliably revealed limitations for adults with DLI (Isaki et al., 2008; Poll et al., 2010). Complex span tasks have revealed differences between adults with DLI and those with typical language, but have been criticized for being another measure of linguistic ability rather than a measure of a memory construct separable from language (Mainela-Arnold & Evans, 2005). Running Span involves memory for digits, which is less dependent on language knowledge. The results from this study, using the Running Span task, support the validity of a somewhat different construct for WM, the capacity of the focus of attention (Cowan et al., 2005). The fact that the task has no separate processing element indicates that the ability of adults with DLI to store information relevant to an immediate cognitive task is more limited than in peers with typical language.

Individual differences in WM capacity had a larger overall effect on sentence repetition in long sentences and in adjunct conditions. The absence of ceiling effects in long sentence conditions may have contributed to the WM × Length interaction. Particularly in short argument conditions, participants with typical language had a limited range of sentence repetition accuracy—all highly accurate. As a result, there was less differentiation across the range of WM capacities. There was a broader range of sentence repetition accuracy in long conditions. This suggests that longer sentences exceeding the WM capacity of most participants are more suited to revealing WM capacity differences.

The WM × Argument Condition interaction suggests that arguments and adjuncts differ in how they stress WM capacity. Adjunct conditions were more demanding of WM capacity, consistent with findings in sentence comprehension studies that arguments ease processing (Boland & Blodgett, 2006; Schutze & Gibson, 1999). Our finding suggests that encountering a word activates its arguments, easing the demand for WM resources in sentence production. Adjuncts may be more demanding because they lack this assist from lexical activation. Adjuncts also have lower co-occurrence frequency than arguments (Schutze & Gibson, 1999). The WM × Argument interaction suggests a refinement to the predictions of the PDH. Instead of a general vulnerability of individuals with DLI to syntactic processing reliant on procedural memory, the difficulty with adjunct processing may be true for individuals with a particular characteristic of many with DLI, limited verbal WM.

Prior work from the perspective of the Baddeley (2000) conception of WM has focused on the role of the episodic buffer in sentence repetition (Baddeley et al., 2009). A critical finding has been that the binding process that creates memory chunks during sentence repetition is not demanding of processing resources (Allen & Baddeley, 2009). Our results are consistent with these findings in that the correlations between sentence repetition and the storage and processing measure (CLPT) and the storage-focused measure (Running Span) were nearly identical. Poorer sentence repetition accuracy for adults with DLI results primarily from limited storage rather than processing.

Limitations

Our language ability groups differed in PIQ and in years of education. Other studies of adults with DLI have also found that adults with DLI score lower on tests of PIQ (Fidler et al., 2011; Lee & Tomblin, 2012; Rost & McGregor, 2012) and have lower educational attainment (Conti-Ramsden & Durkin, 2012). This suggests that our group with DLI was representative of the population of adults with DLI; however, the group with typical language had several members with higher than typical PIQ. PIQ was correlated with processing speed measures, so the unusually high PIQ of some group members with typical language may have extended the range of performance on processing speed tasks, and affected how representative our findings are for adults with typical language. We mitigated effects of our group differences by including education as a covariate in our main analyses. For the group differences in PIQ, we conducted a secondary regression analysis that indicated that language was the larger contributor to variation in sentence imitation when compared with PIQ. We also analyzed effects of processing speed and WM without relying on language ability group classification.

Two measures used in this study, the CLPT and the Truth Value Judgment task, were designed for children. We did not modify the CLPT because we did not observe ceiling effects with pilot participants, but we found that three members of the typical group reached the maximum score for the CLPT. This had the effect of limiting the range of scores in the group with typical language. The simplicity of the Truth Value Judgment task may have also limited our ability to find significant correlations with this measure of processing speed.

Some of the sentences in our sentence repetition task are rather unnatural. We did, however, ensure that sentence plausibility did not differ across conditions, so any effects of the less plausible constructions resulted in a general effect on sentence repetition accuracy rather than a differentiation between conditions.

Clinical and Research Implications

Our findings add to the evidence that sentence repetition differentiates adults with DLI from those with typical language (Poll et al., 2010; Tomblin et al., 1992). The effect size of the group difference varied by length and condition, with long and adjunct condition sentences resulting in larger differences than short and argument conditions. These findings suggest that using longer, adjunct-laden sentences may improve the diagnostic accuracy of sentence repetition tasks for adults, but further testing is required to validate this finding, and to assess diagnostic accuracy. Sentence repetition can be a valuable tool for assessment because it can be used to elicit forms speakers may avoid in spontaneous speech, while also controlling sentences to focus on structures of interest (Seeff-Gabriel, Chiat, & Dodd, 2010). For example, our task enabled us to evaluate predictions of the PDH by varying argument status of sentence constituents while controlling for sentence length, word frequency, and plausibility differences.

Our findings also indicate that more diagnostically accurate sentence repetition tasks result from interactions of participant characteristics and sentence structure. Prior work has shown that language ability status (language impaired or not) interacts with sentence complexity, particularly for sentences with long-distance dependencies such as object relative clauses (Riches et al., 2010). Our findings add that WM and processing speed limitations interact with sentence length and argument status. Adults with DLI and slower processing had more difficulty with long argument sentences that may require more mental operations in the speech production process. Adults with DLI and more limited WM were particularly challenged by longer sentences, and sentences including more adjunct phrases. An established view of sentence repetition performance is that it reflects the language system of the participant (Potter & Lombardi, 1998; Riches, 2012) while also being a measure of WM (Jefferies et al., 2004). Our findings add that sentence repetition accuracy results from cognitive profiles of participants. Limitations in WM or processing speed result in greater vulnerability to sentence repetition difficulties for particular sentence structures.

Our findings provide limited support for the PDH as an explanation for DLI. We did not find the predicted difference in response to adjuncts across language ability groups. We did, however, find that limitations in WM result in particular difficulties with adjunct as compared with argument condition sentences. Eighty-one percent (17 of 21) of participants in the group with DLI were below the study median on WM. Our findings therefore suggest that many adults with DLI will have particular difficulty with adjuncts, but more as a result of their concomitant WM limitation rather than due to their poorer syntactic ability.

Conclusions and Future Directions

In sum, our findings provide evidence for the role of interactions of language and cognitive abilities with sentence structure in sentence repetition accuracy. We found that many study participants with DLI also had WM limitations, which interacted with argument status. This suggests that the difficulties adults with DLI have with certain kinds of grammatical processing may result from WM capacity limitations rather than deficits in learning and processing syntax —based in procedural memory—as predicted by the PDH. Tests of the PDH have focused on assessment of procedural memory, sometimes with correlational links to grammatical ability (Tomblin et al., 2007). The PDH suggests that individuals with DLI may have procedural memory deficits that result in both syntactic difficulties as well as WM limitations. The contribution of this study is to better define which of these traits is at work for one aspect of language processing important to sentence production, processing adjuncts, and arguments.

Acknowledgments

This study was supported by the National Institute on Deafness and Other Communication Disorders under Ruth L. Kirschstein National Research Service Award 1F31DC010960 (awarded to the first author). The views expressed are those of the authors and do not necessarily reflect any official position of the National Institutes of Health.

The authors thank Marlea O'Brien, Connie Ferguson, Marcia St. Clair, and Bruce Tomblin for their generous help with participant recruiting. We also thank the adults who participated in the study.

References

Allen, R. J., & Baddeley, A. D. (2009). Working memory and sentence recall. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 63–85). New York, NY: Psychology Press.

Alloway, T. P., & Gathercole, S. E. (2005). Working memory and short-term sentence recall in young children. *European Journal of Cognitive Psychology, 17,* 207–220. doi:10.1080/09541440440000005

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87,* 85–106. doi:10.1016/j.jecp.2003.10.002

Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research, 52,* 899–914. doi:10.1044/1092-4388(2009/08-0099)

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62,* 67–82.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY: Cambridge University Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4,* 417–423.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language, 61,* 438–456. doi:10.1016/j.jml.2009.05.004

Bishop, D. V. M. (1994). Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics, 15,* 507–550.

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In E. E. Tarone, S. M. Gass, & A. D. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 245–261). Hillsdale, NJ: Erlbaum.

Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer (Version 4.4.30) [Computer software]. Amsterdam: University of Amsterdam. Retrieved from www.praat.org

Boland, J. E., & Blodgett, A. (2006). Argument status and PP-attachment. *Journal of Psycholinguistic Research, 35,* 385–403.

Boland, J. E., & Boehm-Jernigan, H. (1998). Lexical constraints and prepositional phrase attachment. *Journal of Memory and Language, 39,* 684–719.

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 46,* 317–326.

Chomsky, N. (1970). Remarks on nominalization. In R. A. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 184–221). Waltham, MA: Ginn.

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders: A follow-up in later adult life. Cognitive, language, and psychosocial outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 46,* 128–149.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 741–748.

Conti-Ramsden, G., & Durkin, K. (2012). Postschool educational and employment experiences of people with specific language impairment. *Language, Speech, and Hearing Services in Schools, 43,* 507–520. doi:10.1044/0161-1461(2012/11-0067)

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51,* 42–100.

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+). *International Journal of Corpus Linguistics, 14,* 159–190.

de Jong, J., & Fletcher, P. J. (2014). Argument structure and specific language impairment: Retrospect and prospect. In J. Hoeksema & D. Gilbers (Eds.), *Black book: A Festschrift for Frans Zwarts* (pp. 218–228). Groningen, the Netherlands: University of Groningen.

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society, 15,* 331–343. doi:10.1017/S1355617709090481

de Renzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex, 14,* 41–49.

Devescovi, A., & Caselli, M. C. (2007). Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language & Communication Disorders, 42,* 187–208. doi:10.1080/13682820601030686

Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Brookes.

Ellis Weismer, S., Evans, J. L., & 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,* 1249–1260.

Fidler, L. J., Plante, E., & Vance, R. (2011). Identification of adults with developmental language impairments. *American Journal of Speech-Language Pathology, 20,* 2–13. doi:10.1044/1058-0360(2010/09-0096) (Erratum published July, 2013, *American Journal of Speech Language Pathology, 22*, 577)

Fletcher, P., & Garman, M. (1988). Normal language development and language impairment: Syntax and beyond. *Clinical Linguistics & Phonetics, 2,* 97–113.

Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition, 6,* 35–53.

Francis, W., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, MA: Houghton Mifflin.

Gahl, S., Jurafsky, D., & Roland, D. (2004). Verb subcategorization frequencies: American English corpus data, methodological studies and cross-corpus comparisons. *Behavioral Research Methods, 36,* 432–443.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29,* 336–360.

Gaulin, C. A., & Campbell, T. F. (1994). Procedure for assessing verbal working memory in school-age children: Some preliminary data. *Perceptual and Motor Skills, 79,* 55–64.

Grela, B. G., & Leonard, L. B. (2000). The influence of argument-structure complexity on the use of auxiliary verbs by children with SLI. *Journal of Speech, Language, and Hearing Research, 43,* 1115–1125.

Hammill, D. D., & Newcomer, P. L. (1997). *Test of Language Development–Intermediate, Third Edition*. Austin, TX: Pro-Ed.

Hesketh, A., & Conti-Ramsden, G. (2013). Memory and language in middle childhood in individuals with a history of specific language impairment. *PLoS One, 8,* 1–7.

Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson Higher Education.

Hsu, H. J., & Bishop, D. V. M. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Developmental Science, 17,* 352–365. doi:10.1111/desc.12125

Isaki, E., Spaulding, T. J., & Plante, E. (2008). Contributions of language and memory demands to verbal memory performance in language-learning disabilities. *Journal of Communication Disorders, 41,* 512–530. doi:10.1016/j.jcomdis.2008.03.006

Jefferies, E., Ralph, M. A. L., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language, 51,* 623–643. doi:10.1016/j.jml.2004.07.005

Joanisse, M. F., & Seidenberg, M. S. (1998). Specific language impairment: A deficit in grammar or processing? *Trends in Cognitive Science, 2,* 240–247.

Johnson, C. J., Beitchman, J. H., Young, A. R., Escobar, M., Atkinson, L., Wilson, B., . . . Wang, M. (1999). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, and Hearing Research, 42,* 744–760.

Johnston, J., & Kamhi, A. G. (1984). Syntactic and semantic aspects of the utterances of language-impaired children: The same can be less. *Merrill-Palmer Quarterly, 30,* 65–86.

Jones, G., Tamburelli, M., Watson, S. E., Gobet, F., & Pine, J. M. (2010). Lexicality and frequency in specific language impairment: Accuracy and error data from two nonword repetition tests. *Journal of Speech, Language, and Hearing Research, 53,* 1642–1655.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech and Hearing Research, 37,* 418–421.

Kemeny, F., & Lukacs, A. (2010). Impaired procedural learning in language impairment: Results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology, 32,* 249–258. doi:10.1080/13803390902971131

Klem, M., Melby-Lervag, M., Hagtvet, B., Lyster, S. H., Gustafsson, J., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science, 18,* 146–154. doi:10.1111/desc.12202

Lahey, M., Edwards, J., & Munson, B. (2001). Is processing speed related to severity of language impairment? *Journal of Speech, Language, and Hearing Research, 44,* 1354–1361. doi:10.1044/1092-4388(2001/105)

Lee, J. C., & Tomblin, J. B. (2012). Reinforcement learning in young adults with developmental language impairment. *Brain and Language, 123,* 154–163. doi:10.1016/j.bandl.2012.07.009

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

Leonard, L. B., Miller, C. A., & Finneran, D. A. (2009). Grammatical morpheme effects on sentence processing by school-aged adolescents with specific language impairment. *Language and Cognitive Processes, 24,* 450–478. doi:10.1080/01690960802229649

Llorenc, A., Sanz-Torrent, M., Guardia Olmos, J., & MacWhinney, B. (2013). The formulation of argument structure in SLI: An eye-movement study. *Clinical Linguistics & Phonetics, 27,* 111–133. doi:10.3109/02699206.2012.751623

Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short-term memory. *Journal of Memory and Language, 31,* 713–733.

Lum, J. A. G., & Conti-Ramsden, G. (2013). Long-term memory: A review and meta-analysis of studies of declarative and procedural memory in specific language impairment. *Topics in Language Disorders, 33,* 282–297. doi:10.1097/01.TLD.0000437939.01237.6a

Lum, J. A. G., Conti-Ramsden, G., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex, 48,* 1138–1154. doi:0.1016/j.cortex.2011.06.001

Lum, J. A. G., Gelgic, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders, 45,* 96–107. doi:10.3109/13682820902752285

Lust, B., Flynn, S., & Foley, C. (1996). What children know about what they say: Elicited imitation as a research method for assessing children's syntax. In D. McDaniel, C. McKee, & H. S. Cairnes (Eds.), *Methods for assessing children's syntax* (pp. 55–76). Cambridge, MA: MIT Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.

Mainela-Arnold, E., & Evans, J. L. (2005). Beyond capacity limitations: Determinants of word recall performance on verbal working memory span tasks in children with SLI. *Journal of Speech, Language, and Hearing Research, 48,* 897–909.

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,* 416–433.

Miller, C. A., Leonard, L. B., Kail, R., 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,* 712–728.

Miller, C. A., & Poll, G. H. (2009). Response time in adults with a history of language difficulties. *Journal of Communication Disorders, 42,* 365–379. doi:10.1016/j.jcomdis.2009.1004.1001

Morice, R., & McNicol, D. (1985). The comprehension and production of complex syntax in schizophrenia. *Cortex, 21,* 567–580.

Panagos, J. M., & Prelock, P. A. (1982). Phonological constraints on the sentence productions of language-disordered children. *Journal of Speech and Hearing Research, 25,* 171–177.

Pawlowska, M. (2014). Evaluation of three proposed markers for language impairment in English: A meta-analysis of diagnostic accuracy studies. *Journal of Speech, Language, and Hearing Research, 57,* 2261–2273. doi:10.1044/2014_JSLHR-L-13-0189

Poll, G. H., Betz, S. K., & Miller, C. A. (2010). Identification of clinical markers of specific language impairment in adults. *Journal of Speech, Language, and Hearing Research, 53,* 414–429. doi:10.1044/1092-4388(2009/08-0016)

Poll, G. H., Miller, C. A., Mainela-Arnold, E., Donnelly Adams, K., Misra, M., & Park, J. S. (2013). Effects of children's working memory capacity and processing speed on their sentence imitation performance. *International Journal of Language & Communication Disorders, 48,* 329–342. doi:10.1111/1460-6984.12014

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language, 38,* 265–282.

Psychology Software Tools. (2009). E-Prime (Version 2) [Computer Software]. Pittsburgh, PA: Author.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York, NY: Longman Group Limited.

Radford, A. (2004). *Minimalist syntax*. New York, NY: Cambridge University Press.

Redmond, S. M. (2005). Differentiating SLI from ADHD using children's sentence recall and production of past tense morphology. *Clinical Linguistics & Phonetics, 19,* 109–127.

Redmond, S. M., Thompson, H. L., & Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention deficit/hyperactivity disorder. *Journal of Speech, Language, and Hearing Research, 54,* 99–117. doi:10.1044/1092-4388(2010/10-0010)

Riccio, C. A., Cash, D. L., & Cohen, M. J. (2007). Learning and memory performance of children with specific language impairment (SLI). *Applied Neuropsychology, 14,* 255–261.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research, 39,* 1239–1257.

Riches, N. G. (2012). Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language & Communication Disorders, 47,* 499–510. doi:10.1111/j.1460-6984.2012.00158.x

Riches, N. G., Loucase, T., Baird, G., Charman, T., & Simonoff, E. (2010). Sentence repetition in adolescents with specific language impairments and autism: An investigation of complex syntax. *International Journal of Language & Communication Disorders, 45,* 47–60. doi:10.3109/13682820802647676

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57,* 348–379. doi:10.1016/j.jml.2007.03.002

Rost, G. C., & McGregor, K. K. (2012). Miranda rights comprehension in young adults with specific language impairment. *American Journal of Speech-Language Pathology, 21,* 101–108. doi:10.1044/1058-0360(2011/10-0094)

Sattler, J. M., & Ryan, J. J. (1999). *Assessment of children: Revised and updated third edition WAIS-III supplement*. La Mesa, CA: Jerome M. Sattler.

Schuele, C. M., & Tolbert, L. (2001). Omission of obligatory relative markers in children with specific language impairment. *Clinical Linguistics & Phonetics, 15,* 257–274.

Schutze, C. T., & Gibson, E. (1999). Argumenthood and English prepositional phrase attachment. *Journal of Memory and Language, 40,* 409–431.

Seeff-Gabriel, B., Chiat, S., & Dodd, B. (2010). Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties. *International Journal of Language & Communication Disorders, 45,* 691–702. doi:10.3109/13682820903509432

Semel, E., Wiig, E., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals, Fourth Edition*. San Antonio, TX: The Psychological Corporation.

Sitlington, P. L. (2003). Postsecondary education: The other transition. *Exceptionality, 11,* 103–113.

Slobin, D. E., & Welch, C. A. (1971). Elicited imitation as a research tool in developmental psycholinguistics. In C. S. Lavatelli (Ed.), *Language training in early childhood education* (pp. 170–185). Urbana, IL: University of Illinois Press for the ERIC Clearinghouse.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys and humans. *Psychological Review, 99,* 195–231.

Stokes, S. F., Wong, A. M., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research, 49,* 219–236.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education, Inc.

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. (2014). Evidence for distinct cognitive profiles in autism spectrum disorders and specific language impairment. *Journal of Autism and Developmental Disorders, 44,* 19–30. doi:10.1007/s10803-013-1847-2

Thompson, S. A. (1997). Discourse motivations for the core-oblique distinction as a language universal. In A. Kamio (Ed.), *Directions in functional linguistics* (pp. 59–82). Philadelphia, PA: John Benjamins.

Thordardottir, E. T., & Ellis Weismer, S. (2002). Verb argument structure weakness in specific language impairment in relation to age and utterance length. *Clinical Linguistics & Phonetics, 16,* 233–250.

Tomblin, J. B., Freese, P. R., & Records, N. L. (1992). Diagnosing specific language impairment in adults for the purpose of pedigree analysis. *Journal of Speech and Hearing Research, 35,* 832–844.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development, 3,* 269–293.

Tomblin, J. B., Zhang, X., Buckwalter, P., & O'Brien, M. (2003). The stability of primary language disorder: Four years after kindergarten diagnosis. *Journal of Speech, Language, and Hearing Research, 46,* 1283–1296. doi:10.1044/1092-4388(2003/100)

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385–398.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30,* 37–69.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41,* 399–433.

van Gompel, R. P. G., & Pickering, M. J. (2007). Syntactic parsing. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 289–307). Oxford, England: Oxford University Press.

Vang Christensen, R., & Hansson, K. (2012). The use and productivity of past tense morphology in specific language impairment: An examination of Danish. *Journal of Speech, Language, and Hearing Research, 55,* 1671–1689. doi:10.1044/1092-4388(2012/10-0350)

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale–Third Edition.* San Antonio, TX: The Psychological Corporation.