JSLHR

Research Article

Language Skill Mediates the Relationship Between Language Load and Articulatory Variability in Children With Language and Speech Sound Disorders

Janet Vuolo[a] and Lisa Goffman[b]

Purpose: The aim of the study was to investigate the relationship between language load and articulatory variability in children with language and speech sound disorders, including childhood apraxia of speech.
Method: Forty-six children, ages 48–92 months, participated in the current study, including children with speech sound disorder, developmental language disorder (aka specific language impairment), childhood apraxia of speech, and typical development. Children imitated (low language load task) then retrieved (high language load task) agent + action phrases. Articulatory variability was quantified using speech kinematics. We assessed language status and speech status (typical vs. impaired) in relation to articulatory variability.
Results: All children showed increased articulatory variability in the retrieval task compared with the imitation task. However, only children with language impairment showed a disproportionate increase in articulatory variability in the retrieval task relative to peers with typical language skills.
Conclusion: Higher-level language processes affect lower-level speech motor control processes, and this relationship appears to be more strongly mediated by language than speech skill.

Language and speech motor control develop jointly in children who are acquiring speech and language typically as well as those with delayed or disordered communication. Though obviously codeveloping systems, language, speech, and speech motor control have usually been investigated separately. The relationship between language, speech, and speech motor control is complex and bidirectional, and systematic investigation into the nature of these interactive processes has the potential to further our understanding of typical as well as disordered speech and language acquisition. The aim of the current study was to examine how language load affects articulatory control in children with typical and impaired language and/or speech skills.

[a]Department of Speech and Hearing Science, The Ohio State University, Columbus
[b]Callier Center for Communication Disorders, Behavioral and Brain Sciences, University of Texas at Dallas
Correspondence to Janet Vuolo: vuolo.4@osu.edu

Editor-in-Chief: Sean Redmond
Editor: Lisa Archibald
Received February 12, 2018
Revision received June 27, 2018
Accepted July 11, 2018
https://doi.org/10.1044/2018_JSLHR-L-18-0055

The Development of Speech Motor Control

Speech motor control has been investigated using a variety of measures, from those that index individual movement parameters (e.g., duration, displacement, or peak velocity at a single time point) to composite measures that incorporate spatial and temporal features of a multimovement sequence. The spatiotemporal index (STI; Smith, Goffman, Zelaznik, Ying, & McGillem, 1995) has been used to index the trial-to-trial variability of multimovement sequences. The STI quantifies the degree to which a speaker produces a syllable, word, phrase, or sentence with the same spatial and temporal patterning of the upper lip, lower lip, and jaw across multiple trials.

Shifts in articulatory variability are indicative of typical maturational processes as well as disordered articulatory control. In typically developing (TD) children, an expressive vocabulary spurt observed at 24 months of age is associated with a temporary increase in jaw movement variability, presumably as the speech motor system reorganizes to accommodate rapid cognitive and linguistic development (Green, Moore, & Reilly, 2002; Nip, Green, & Marx, 2011). As a group, children produce more variable

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

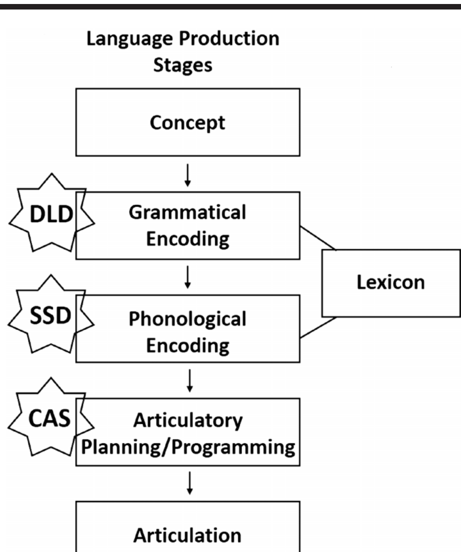
articulatory movements than adults (Goffman & Smith, 1999; Maner, Smith, & Grayson, 2000; Smith & Goffman, 1998; Smith & Zelaznik, 2004). Articulatory variability does not decrease to adultlike levels until late adolescence (Smith & Zelaznik, 2004; Walsh & Smith, 2002). Maturation results in more stable and precise spatial and temporal coupling of the lips and jaw (Green et al., 2002; Smith & Zelaznik, 2004). The stabilization of the speech motor system is nonlinear and nonuniform (Nittrouer, 1993; Smith & Goffman, 1998) and is tied to the development of other cognitive, linguistic, and perceptual processes (Goffman, 2010; Nip et al., 2011).

High levels of articulatory variability may also reflect impaired speech motor control, particularly in persons with motor speech disorders such as childhood apraxia of speech (CAS) and apraxia of speech, the adult analog (e.g., Grigos, Moss, & Lu, 2015). Motor planning or programming deficits may result in the variable realization of spatial and temporal speech movements (American Speech-Language-Hearing Association [ASHA], 2007), which can be detected through speech kinematics. Articulatory variability measures such as the STI are well suited to apply to what are presumably relatively low-level motor disorders such as CAS.

Language Load: Imitation Versus Retrieval

Figure 1 shows a simplified hierarchical model of the processing levels involved in language production (adapted from Levelt, Roelofs, & Meyer, 1999), with language providing top-down input into articulation. In contrast, speech motor control researchers have generally focused almost exclusively on articulatory planning/programming and

Figure 1. Language production model (adapted from Levelt, Roelofs, & Meyer, 1999), illustrating presumed primary level of impairment in developmental language disorder (DLD), speech sound disorder (SSD), and childhood apraxia of speech (CAS).



articulatory implementation levels. This divide has resulted in a long-standing perception that speech motor control operates independently from language processes. However, more recent experimental studies have documented bidirectional interactions across grammatical, lexical, phonological, and phonetic levels (e.g., Goldrick, Baker, Murphy, & Baese-Berk, 2011). For example, syntactic complexity (Kleinow & Smith, 2000; Sadagopan & Smith, 2008; Saletta, Goffman, Ward, & Oleson, 2018), lexical status (Heisler & Goffman, 2016; Heisler, Goffman, & Younger, 2010), and prosodic complexity (Goffman, 1999, 2004) affect articulatory movement patterning in typical and atypical speakers.

In the current study, we focus on the relationship between language load (i.e., the relative language demands associated with the speaking task) and articulatory variability. Language load can be operationalized in multiple ways. We focus on imitation versus retrieval cues, as these are commonly used in studies of adult language processing (e.g., Dell, Martin, & Schwartz, 2007) and as intervention approaches (e.g., Strand, Stoeckel, & Baas, 2006; van der Merwe, 2011). However, very few studies of speech motor control incorporate active retrieval of language. Systematically manipulating language load in kinematic studies will enhance our understanding of the relationship between language and articulatory processes in children with typical and atypical communication skills.

Referring again to Figure 1, imitation tasks are thought to primarily invoke language processing from the phonological encoding level down through articulatory output (Dell et al., 2007). Assuming the items to be imitated do not exceed working memory capacity, the participant is able to chunk the information and repeat it without engaging in robust language processing (e.g., Baddeley, Hitch, & Allen, 2009; Dell et al., 2007). In contrast to imitation, retrieval requires additional cognitive and language demands such as perceptual, semantic, and grammatical processing (Kurland, Reber, & Stokes, 2014), as well as phonological processing and articulatory planning and execution. In other words, retrieval invokes every language production level depicted in Figure 1. Whereas imitation can be viewed as a rote task with relatively shallow language processing demands, retrieval is a relatively high language load task that necessitates active language generation.

Retrieval also relies on procedural memory processes to a greater extent than imitation. The procedural memory system is responsible for the acquisition, storage, and retrieval of new cognitive and motor skills (Gabrieli, 1998). There is some evidence that, for children with developmental language disorder (DLD; also referred to as specific language impairment), aspects of language that rely on the procedural memory system are impaired whereas those that rely on the declarative memory system are relatively unaffected (Hsu & Bishop, 2014; Saletta et al., 2018; Ullman & Pullman, 2015). For example, the procedural memory system supports syntactic, morphological, and phonological rule learning (Lum, Conti-Ramsden, Page, & Ullman, 2012), all of which are affected in DLD (Ullman & Pierpont, 2005).

In contrast, the declarative memory system supports explicit knowledge about facts and events, as well as semantic and lexical knowledge; these domains appear to be relatively intact in children with DLD (Ullman & Pullman, 2015). Therefore, a reasonable prediction is that procedural memory deficits in children with DLD result in greater difficulty with language formulation compared with rote imitation.

In an exploratory study including nine children with and without speech disorders, we demonstrated the feasibility of using imitation and retrieval tasks to explore how language load affects articulatory variability (Vuolo & Goffman, 2017). We found that all children produced agent + action phrases in a retrieval context with higher articulatory variability compared to imitation, irrespective of diagnostic category. Articulatory variability can be used to index language load: Higher language processing demands disrupt the stability and patterning of articulatory movements. As an extension of this work, in the current study we suggest that language load should differentiate some diagnostic categories from others. Specifically, articulatory variability measures should discriminate children with language impairment (i.e., children with DLD and those children with CAS and concomitant language impairment) in the high language load retrieval task and discriminate children with low-level motor planning, programming, or implementation impairment (i.e., children with CAS or speech sound disorder [SSD]) in the low language load imitation task.

Children with DLD show a primary impairment in higher-level language processes such as the acquisition of grammar (see Figure 1; e.g., Leonard, 2014). Several studies have also documented higher articulatory variability in children with DLD compared with TD peers (Brumbach & Goffman, 2014; Goffman, 1999, 2004; Heisler et al., 2010). The articulatory deficits observed in DLD may be due to the higher-level language impairments that characterize this population. Alternatively, children with DLD may have comorbid impairments at lower-level motor stages. As a group, children with DLD show limb motor deficits across a range of motor tasks (Brumbach & Goffman, 2014; Hill, 2001; McPhillips, Finlay, Bejerot, & Hanley, 2014; Zelaznik & Goffman, 2010), but these deficits may pattern with the specific procedural memory mechanisms invoked during certain tasks, such as those which require sequencing and coordination (Hsu & Bishop, 2014; Sanjeevan & Mainela-Arnold, 2017; Vuolo, Goffman, & Zelaznik, 2017). The relationship between limb, language, and articulatory deficits in DLD remains unclear.

Children with CAS are presumed to have a motor speech disorder that is characterized by deficits in planning or programming the spatial and temporal parameters of speech movement sequences (see Figure 1; ASHA, 2007, p. 4). In other words, the crux of the speech deficit resides not at a linguistic or phonological processing level (such as is supposed for children with DLD and SSD, respectively) but at the level of articulatory planning and control. Children with CAS show impairments in the precision and consistency of speech movements (ASHA, 2007) and therefore should show high levels of articulatory variability even in tasks with relatively low language demands. The few kinematic studies focused on children with CAS have documented higher articulatory variability in this population compared with nonapraxic peers (Case & Grigos, 2016; Grigos et al., 2015; Moss & Grigos, 2012; Terband, Maassen, van Lieshout, & Nijland, 2011). However, in all but the Terband et al. (2011) study, the relative language demands associated with the experimental tasks were high, requiring children to retrieve mono- and multisyllabic words and/or novel words. Because most children with CAS also show language impairments (Lewis et al., 2004), it is not clear that increased articulatory variability in children with CAS can be attributed to motor speech and not to language difficulties.

Duration measures may also distinguish children with CAS from peers without CAS. Children with CAS may reduce their speaking rate as a compensatory strategy to allow more time to plan/program speech movement sequences (Case & Grigos, 2016; Grigos et al., 2015; Grigos & Kolenda, 2010), which could, in turn, result in more stable articulatory movements. However, the few studies that have examined movement durations in children with CAS have reported mixed results. Case and Grigos (2016) found that children with CAS produced longer movement durations compared with TD peers (SSD children were not included), but in an earlier study (Grigos et al., 2015), movement duration did not distinguish children with CAS from those with SSD. The relationship between movement durations and diagnostic category remains open to investigation.

Surprisingly, little is known about articulatory variability in children with SSD (also referred to as developmental phonological disorder). We do not know if children with SSD, due to their speech production deficits, show higher articulatory variability relative to those with typical development. Returning to Figure 1, children with SSD are thought to show a primary deficit in phonology (Rvachew & Brosseau- Lapré, 2016) and do not show language impairments or apraxic speech features. Therefore, including children with SSD allows us to investigate performance across three clinical groups, each of whom are presumed to show a primary deficit at a different stage of the language production model. In addition, because most young children with DLD and CAS also show speech sound impairments (e.g., Alt, Plante, & Creusere, 2004; ASHA, 2007; Deevy, Weil, Leonard, & Goffman, 2010), children with SSD serve as a type of control group. Comparing performance across all three diagnostic categories will allow for specification of the influences of language and motor speech skill on articulatory control processes.

Current Study

The goal of the current study is to assess how language load affects articulatory variability. We include children with typical development, as well as those with diagnoses considered to be lexical-semantic and grammatical (DLD), phonological (SSD), or motor (CAS) in origin. The first objective is to assess how language status and speech status influence articulatory variability. We predict that all children will show increased articulatory variability in the retrieval task, but that children with language impairment, in particular (DLD and those children with CAS + language impairment), will show higher articulatory variability in retrieval compared with peers with typical language skills. The second objective is to begin to address, via a small group of children who meet the diagnostic criteria for CAS, how the presence of a presumed motor speech impairment influences articulatory variability. To assess this, we compare articulatory variability across the four groups. We expect that children with CAS will show high levels of articulatory variability in the imitation task compared with peers without CAS. We also assess the durations of the extracted movement sequences in imitation and retrieval to determine if children with CAS show longer durations (reflecting increased time required to organize articulatory movement sequences) compared with nonapraxic peers. The inclusion of children with SSD serves as an important control to assess the role of language disorder on speech motor control in high and low language load conditions.

Method

Participants

The institutional review board at Purdue University approved all tasks and procedures. Parental consent and child assent were obtained prior to participation. Forty-six children (20 girls, 26 boys) across four groups participated in this project. The four groups included children who were TD (n = 14; nine girls, five boys), children with SSD (n = 14; four girls, 10 boys), children with DLD (n = 12; six girls, six boys), and children with CAS (n = 6; one girl, five boys). Children ranged in age from 48 to 92 months. A one-way analysis of variance (ANOVA) revealed no age differences across the four groups, $F(3, 42) = 1.52$, $p = .22$, $\eta_p^2 = .10$.

All children were recruited by targeting each of the four groups and then confirming group assignment through our inclusionary criteria (specified below). Children with SSD were recruited from local schools and clinics in the Greater Lafayette, Indiana, area by mailing specific recruitment fliers to speech-language pathologists and by posting fliers in the M. D. Steer Speech Clinic at Purdue University. Children with DLD were recruited to participate in the Purdue Summer Fun Preschool Program, a therapy and research program for children with DLD. All of the children participating in this program met criteria for DLD and participated in studies directed by Laurence Leonard and Lisa Goffman. The diagnostic procedures used to verify DLD status are delineated below. It is important to note that the inclusionary and exclusionary procedures are identical to those used in other studies of DLD (e.g., Haebig, Leonard, Usler, Deevy, & Weber, 2018). Children with CAS were recruited through advertisements posted on Apraxia Kids (formerly the Childhood Apraxia of Speech of North America Association). Children with typical development were recruited through research advertisements posted in Purdue Today, a faculty and staff newsletter, and by contacting appropriate families in the Speech, Language, and Hearing Science research database.

Children attended three individual sessions, each lasting 1–1.5 hr. In the first session, the child completed the imitation and retrieval tasks and began the developmental testing. In the remaining two sessions, the child completed two additional experimental tasks, not reported here, and finished the remaining developmental testing. Parents were compensated $10 per each hour of participation to defer travel costs, and children chose a small toy at the end of each session.

Inclusionary Criteria for All Participants

All children were monolingual English speakers with no history of neurological problems (e.g., seizures, epilepsy, head injury), no previous diagnosis of autism spectrum disorder, and normal or corrected-to-normal vision per parent report. In addition, all children passed a bilateral hearing screening (20 dB HL at 500, 1000, 2000, and 4000 Hz), scored in the minimal to no symptoms of autism spectrum disorder range on the Childhood Autism Rating Scale–Second Edition (Schopler, Reichler, & Renner, 2010), and received a structural oral mechanism score within the normal range on the Robbins and Klee (1987) oral motor protocol. All children completed the Primary Test of Nonverbal Intelligence (Ehrler & McGhee, 2008). A one-way ANOVA revealed no group differences in nonverbal reasoning ability, $F(3, 42) = 0.49$, $p = .69$, $\eta_p^2 = .14$. Table 1 summarizes the group assessment data.

Standardized Assessment Battery

All children completed a battery of language and speech assessments (described below and summarized in Table 1) to determine language status, speech status, and group assignment.

Language Measures

Children were classified as DLD in the current study based on Leonard (2014). This included a standard score below 87 on the Structured Photographic Expressive Language Test–Preschool: Second Edition (SPELT-P 2; Dawson et al., 2005; n = 10) or a standard score of 89 and a developmental sentence score (Lee & Canter, 1971) below the 10th percentile (n = 2). A cutoff of 87 on the SPELT-P 2 shows high sensitivity and specificity for this age group (Greenslade, Plante, & Vance, 2009). As a group, children classified as DLD scored below the average range on the SPELT-P 2 (M = 79.9, SD = 6.9).

Four children who were older than the 5;11 (years; months) cutoff on the SPELT-P 2 (for TD, n = 1; for SSD,

Table 1. Group characteristics.

Measure	TD (n = 14)		SSD (n = 14)		DLD (n = 12)		CAS (n = 6)	
	M (*SD*)	Range	*M* (*SD*)	Range	*M* (*SD*)	Range	*M* (*SD*)	Range
Age in months	62.9 (8.9)	48–75	58.6 (9.7)	50–72	58.8 (4.1)	52–65	69.0 (17.9)	50–92
PTONI	118.0 (17.9)	89–149	113.1 (16.3)	80–149	112.2 (17.4)	89–139	108.7 (18.9)	84–140
SPELT-P 2/SPELT-3	110.5 (7.4)	99–125	99.1 (10.8)	80–117	79.9 (6.9)	68–89	66.0 (24.5)	40–111
BBTOP	99.3 (6.2)	88–107	69.0 (5.7)	65–83	76.3 (12.1)	65–101	66.7 (3.6)	65–74
DEAP WI	10.0 (8.3)	0–28	20.9 (9.9)	4–32	21.1 (9.1)	8–36	49.3 (11.1)	32–62
CAS checklist	0.3 (0.6)	0–2	1.1 (0.7)	0–2	1.0 (0.7)	0–2	5.7 (1.2)	4–7
TSS	23.5 (0.5)	23–24	23.0 (0.8)	22–24	23.8 (0.6)	22–24	23.5 (0.5)	23–24
TFS	108.6 (3.0)	104–112	105.3 (4.1)	98–112	104.6 (6.4)	92–111	99.5 (8.7)	86–108
DDK	1.5 (0.7)	0–2	0.9 (0.8)	0–2	0.7 (0.8)	0–2	0.5 (0.8)	0–2
CARS-2	15.5 (0.7)	15–16.5	15.4 (0.8)	15–18	16.1 (1.1)	15–18.5	15.7 (0.7)	15–16.5

Note. TD = typically developing; SSD = speech sound disorder; DLD = developmental language disorder; CAS = childhood apraxia of speech; PTONI = Primary Test of Nonverbal Intelligence; SPELT-P 2 = Structured Photographic Expressive Language Test–Preschool 2; SPELT-3 = Structured Photographic Expressive Language Test–Third Edition; BBTOP = Bankson–Bernthal Test of Phonology; DEAP WI = Diagnostic Evaluation of Articulation and Phonology Word Inconsistency subtest; TSS = total structural score (Robbins & Klee, 1987); TFS = total functional score (Robbins & Klee, 1987); DDK = diadochokinetic rate score (Robbins & Klee, 1987); CARS-2 = Childhood Autism Rating Scale–Second Edition.

n = 1; for CAS, n = 2) were administered the Structured Photographic Expressive Language Test–Third Edition (SPELT-3; Dawson et al., 2003). Children in the TD group received a standard score within the average range on the SPELT-P 2/SPELT-3 (M = 110.5, SD = 7.4). Similarly, children with SSD did not show evidence of expressive language impairment based on performance on the SPELT-P 2/SPELT-3 (M = 99.1, SD = 10.8). However, one child with SSD received a SPELT-P 2 score of 80; this child demonstrated pervasive final consonant deletion, and we determined that her speech sound deficit influenced her production of final position grammatical morphemes. Because most children with CAS also show impaired language skills (Lewis et al., 2004), expressive language performance was free to vary for children in the CAS group. Five out of six children with CAS scored below the average range on the SPELT-P 2/SPELT-3 (M = 66.0, SD = 24.5).

Speech Measures

The Bankson–Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990) is a standardized articulation assessment that scores consonant production accuracy in initial and final word positions as children name familiar words. All TD children scored within the average range (M = 99.3, SD = 6.2). Children with SSD (M = 69.0, SD = 5.7) and CAS (M = 66.7, SD = 3.6) were required to score at least 1 *SD* below the mean. Performance in the DLD group was free to vary. Nine out of 12 children with DLD showed speech impairments on the BBTOP (M = 79.9, SD = 6.9).

The Diagnostic Evaluation of Articulation and Phonology Word Inconsistency subtest (Dodd, Hua, Crosbie, Holm, & Ozanne, 2006) is the only published measure that scores word inconsistency the same way that ASHA (2007) defines inconsistency in CAS. Children name 25 pictures across three separate trials, and the number of productions that are transcribed differently using broad phonemic transcription is divided by 25. An inconsistency score of 40% or greater is used to classify CAS and inconsistent phonological disorder. Variation between developmentally appropriate errors and the correct production are not scored as inconsistent. Children who received a score of 40% inconsistent or greater in conjunction with impaired oral motor skills were classified as CAS (M = 49.3, SD = 11.1). One participant in the CAS group (age 7;6) showed clear CAS speech features (seven features on the CAS checklist, described below) but received a word inconsistency score of 32%. Children with typical development (M = 10.0, SD = 8.3), SSD (M = 20.9, SD = 9.9), and DLD (M = 21.1, SD = 9.1) scored below the 40% cutoff.

The CAS checklist (Shriberg, Potter, & Strand, 2011) is also used to classify CAS. Children with CAS must demonstrate four or more out of 10 speech features, and children without CAS must demonstrate fewer than four speech features. The 10 speech features include vowel distortions, difficulty in achieving initial articulatory configurations or transitionary movement gestures, equal stress or lexical stress errors, distorted substitutions, syllable segregation, groping, intrusive schwa, voicing errors, slow rate, and increased difficulty with multisyllabic words. Children with CAS demonstrated four or more speech features. Children with typical development, SSD, and DLD demonstrated fewer than four speech features associated with CAS. See Table 1 for means and standard deviations by group and Table 2 for the number of participants per group who demonstrate each CAS feature.

The oral and speech motor control protocol (Robbins & Klee, 1987) was administered to all children to assess the structure and function of the oral mechanism. As mentioned previously, all children were required to achieve a total structural score in the average range. The total functional score was free to vary for all children. For the diadochokinetic rate measure, children were asked to repeat /pərəkək/ as many times as possible within 3 s. A score of 0 indicated

Table 2. Number of participants per group demonstrating each childhood apraxia of speech (CAS) feature.

CAS feature	TD	SSD	DLD	CAS
Vowel distortions	0	3	1	5
Initial articulatory configurations/transitionary movement gestures	0	0	0	6
Equal stress/lexical stress errors	0	0	0	0
Distorted substitutions	2	4	4	4
Syllable segregation	0	0	0	2
Groping	2	4	3	5
Intrusive schwa	0	0	0	2
Voicing errors	0	3	2	6
Slow rate	0	0	0	0
Increased difficulty with multisyllabic words	0	1	3	4
Total	4	15	13	34

Note. CAS features are from Shriberg et al. (2011). TD = typically developing; SSD = speech sound disorder; DLD = developmental language disorder.

the child was unable to perform the task, a score of 1 indicated that this skill was emerging, and a score of 2 indicated that the child performed the task accurately. Scores are reported in Table 1.

Materials

Audio stimuli for the imitation task were recorded in a soundproof booth by an adult native English-speaking female talker. The female talker produced the target phrases in a natural, child-friendly voice. Audio files were equated for intensity in Praat (Boersma & Weenink, 2013) and paired with a corresponding video clip (described below) in PowerPoint. High-quality video and audio recordings of the imitation and retrieval tasks and relevant diagnostic testing (e.g., speech and language sample and standardized measures of production) were obtained.

Procedure

Signals Recorded

The 3D Investigator (Northern Digital, Inc.) was used to record lip and jaw movements while the child produced the target phrases in imitation and retrieval. Three infrared light-emitting diodes were attached to the child's face to record lip and jaw movement: one placed at midline on the vermillion border of the upper lip, one at midline on the lower lip, and one on a small splint attached underneath the jaw with medical adhesive. Five additional diodes were used to provide a reference frame and account for head movement: one on the forehead at midline and four placed on a pair of sports goggles worn by the child. Kinematic data were collected at a sampling rate of 250 samples per second. A time-locked acoustic signal was also collected at a sampling rate of 16,000 Hz.

Experimental Tasks

Each child imitated agent + action phrases (imitation task) and then retrieved those phrases (retrieval task) in response to video clips. The child was seated approximately 8 ft from a 30-in. Dell monitor, on which the visual stimuli were displayed. The 3D Investigator movement tracking camera has three fixed lenses and was wall-mounted above the monitor. All tasks were presented from a laptop computer connected to the monitor. Because response times were not required, all stimuli were presented via PowerPoint. This allowed the examiner to present the visual stimuli at a comfortable pace for each child. The audio stimuli were presented at a comfortable loudness level.

Stimuli and presentation. Six video clips, each approximately 2 s in length, depicted puppets performing various actions. Each video corresponded to a phrase, including two target phrases (baby pops and mommy beeps) and four foils (mommy bumps, puppy mops, baby puffs, and puppy wipes). The foils were included so that children were required to generate a variety of phrases in the retrieval task, thereby increasing the language load. These foils were not analyzed. Both target phrases included only labial consonants and thus could be subjected to kinematic analyses. Baby pops and mommy beeps were each paired with a sound effect (i.e., a popping sound and a car horn beeping, respectively) to increase salience and clarify meaning. In both the imitation and retrieval tasks, the two target phrases were presented 12 times each and the four foils were presented four times each to elicit a total of 40 productions. The phrases were presented in quasirandom order (no more than two consecutive presentations of the same target) in four blocks of 10 to allow for rest breaks as needed.

The imitation task was presented first, followed by the retrieval task. This allowed the child to become familiar with the phrases before retrieving them independently. For the imitation task, the examiner instructed the child, "You are going to see some videos of puppets doing different things. A woman will tell you what they are doing. Try to say exactly what the woman says." The corresponding audio and video files were presented simultaneously. For the retrieval task, the child was told, "Now it is your turn to tell me about the videos. Say the same sentences you just practiced."

To minimize loss of data, two prompting procedures were used if the child produced an error. For both the imitation and retrieval tasks, a segmental prompting procedure was used if the child omitted a labial consonant or substituted a nonlabial consonant. The examiner provided up to three gestural phonemic prompts to attempt to elicit the target sound. For example, if a child produced “huppy” for puppy, the examiner paired a gestural cue (touching an index finger to the examiner’s cheek) with the prompt “That starts with a /p/. Say puppy.” The examiner provided these prompts only during the first three presentations of the video.

For the retrieval task, a prompting procedure was used if a participant forgot a target word or used an incorrect word. In this prompting procedure, if the child forgot one word or provided an incorrect word, the examiner provided the correct word and asked the participant to produce the phrase again (e.g., if the child produced “Doggy mops,” the examiner responded, “That’s a puppy. Tell me that again”). If the child forgot both words or produced both incorrectly (e.g., “Doggy sweeps”), the examiner played the audio clip that was used in the imitation task and asked the participant to imitate the phrase. Phrases produced in direct imitation during the retrieval task were excluded from kinematic analyses.

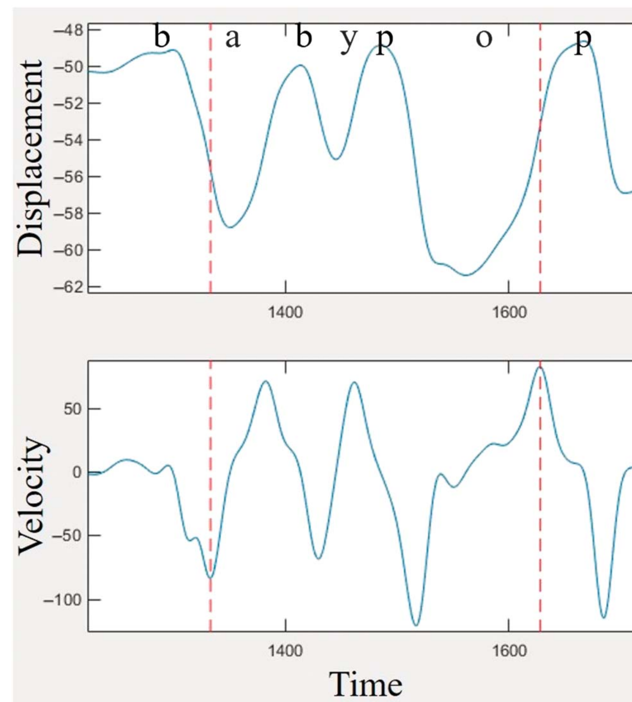
Kinematic Data Processing

Upper lip, lower lip, and jaw signals were recorded, and lip aperture (calculated as the difference between the upper lip and lower lip displacement records) was derived (Smith & Zelaznik, 2004). The movements associated with the phrases were extracted from the long data files in MATLAB (MathWorks, 2009). All stimuli contained only labial consonants (/p/, /b/, /m/, and /f/) and so were amenable to kinematic analysis.

For the imitation and retrieval tasks, the first 10 usable productions were included in the kinematic analyses. Productions that contained missing signals from the lip or jaw diodes, laughter, or disfluencies were excluded from analyses. The minimum number of records extracted for the kinematic analyses was five. As shown in Figure 2, displacement records were extracted by selecting the peak velocity that corresponded to the lip closure at the onset and offset of the target phrase. Movement onsets and offsets were initially selected via visual inspection of the displacement record. An algorithm determined the peak velocity within a 25-point (100-ms) window of the examiner selected point. Synchronized acoustic signals confirmed selections of the kinematic records.

Figure 2 shows an example of how the peak velocity points were selected. The lip opening movement at the beginning of the word marked the onset of the target (e.g., in baby pops, the onset for the initial [b]). Movement offsets corresponded with lip closing movement for the final consonant (e.g., for the word-final [p] in [pops]). Productions were included if the child produced a labial consonant in initial and final positions, even if that consonant was in error. In the retrieval task, extracted productions that were +2 *SD*s of the mean duration were discarded as outliers.

Figure 2. Example of the kinematic data extraction. Lower lip movement from a child with childhood apraxia of speech imitating the phrase baby pops. Red dotted lines show where the velocity algorithm detected the peak velocity of the opening (out of the initial [b] in baby) and closing (into the final [p] in pops) labial movements.



Calculation of Articulatory Variability

Lip aperture (upper lip minus lower lip) was calculated from the first 10 extracted movement sequences. The records were then amplitude- and time-normalized to eliminate differences in amplitude (i.e., loudness) and duration (i.e., speech rate). Following normalization, standard deviations were computed at 2% intervals across all the normalized displacement records. The sum of these 50 standard deviations yields the STI (Smith et al., 1995), which is used to quantify articulatory variability over multiple productions of a syllable, word, or phrase. A higher STI value indicates higher articulatory variability. Figure 3 shows eight imitated productions of the target phrase “baby pops” from a child with CAS, with the nonnormalized records (top panel), normalized records (middle panel), and the STI (bottom panel).

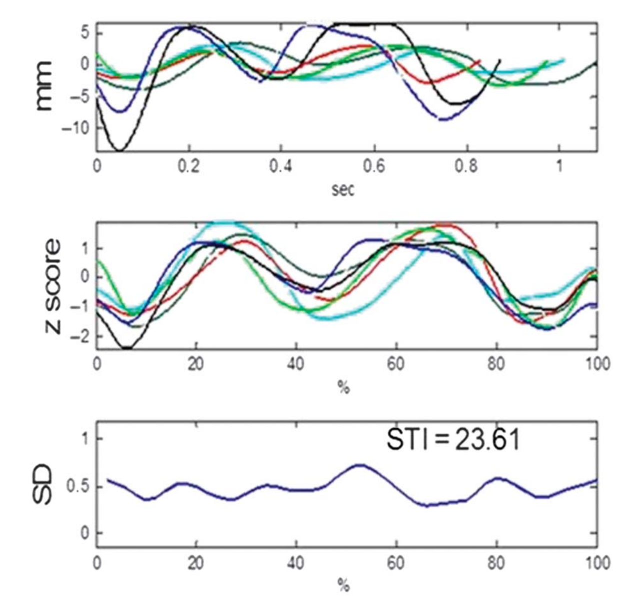
Calculation of Duration

Duration was measured as the time (in seconds) between the movement onset and offset points selected in MATLAB. From these extracted records, means and standard deviations were computed for each phrase in imitation and retrieval.

Analyses

The overall statistical design included mixed ANOVAs. Because the central goal was to determine the role of language

Figure 3. Calculation of the spatiotemporal index (STI). Example of eight imitated productions of the target phrase *baby pops* from a child with childhood apraxia of speech, with the nonnormalized (top) and normalized (middle) records and the STI (bottom).



disorder and speech disorder in speech motor performance, mixed ANOVAs were performed using language status (i.e., typical vs. impaired language performance on the SPELT-P 2/SPELT-3) and speech status (i.e., typical vs. impaired speech performance on the BBTOP) as the between-subjects factor, irrespective of group assignment. An additional analysis examined whether children with CAS showed high levels of articulatory variability in the low-load imitation task compared with children without CAS. In this analysis, more specific group assignment (TD, SSD, DLD, and CAS) served as between-subjects factors. Because of the small number of CAS participants, we relied on group means and effect sizes to interpret these findings.

Within-subject factors included task (imitation and retrieval) and phrase ("baby pops" and "mommy beeps") for the dependent measure STI. In the ANOVA by group, we also included the dependent measure duration. The Tukey's honestly significant difference procedure was used to determine which comparisons were significantly different. An alpha level below .05 was considered significant for all statistics. Effect sizes are reported for all results.

Results

In the imitation task, the mean number of productions included in each STI ranged from 9.58 (CAS group) to 9.96 (TD and SSD groups). In the retrieval task, the mean number of productions included in each STI ranged from 9.32 (TD group) to 9.68 (SSD group). A one-way ANOVA revealed no differences in the number of productions included per group in the analyses for the imitation, $F(3, 88) = 1.58$, $p = .20$, $\eta_p^2 = .05$, or retrieval, $F(3, 88) = 0.49$, $p = .69$, $\eta_p^2 = .02$, tasks.

Articulatory Variability by Language Status

To assess the relationship between language skill and articulatory variability in the imitation and retrieval tasks, children were separated into two groups based on performance on the SPELT-P 2/SPELT-3. Seventeen children showed impaired language ability on the SPELT-P 2/SPELT-3, including all 12 children with DLD and five out of six children with CAS. Fourteen of the 17 children with language impairment also showed speech sound deficits; therefore, it was not possible to exclude children with language impairment and concomitant speech deficits from this analysis. Children were coded as having either language performance in the average range ($n = 29$) or impaired language performance ($n = 17$) on the SPELT-P 2/SPELT-3.

Figure 4 shows the mean STIs by language status in the imitation and retrieval tasks. STI means and standard deviations are also reported in Table 3. There was no main effect of language status, $F(1, 44) = 2.32$, $p = .14$, $\eta_p^2 = .05$. All children showed increased articulatory variability in the retrieval task, $F(1, 44) = 18.68$, $p < .001$, $\eta_p^2 = .30$. However, there was a significant task by language status interaction, $F(1, 44) = 8.03$, $p = .01$, $\eta_p^2 = .15$; children with language impairment demonstrated higher articulatory variability in the retrieval task compared with children with typical language skills. Children with language impairment were especially vulnerable to the increased language load obligated in the retrieval task.

Articulatory Variability by Speech Status

To examine how speech skill relates to articulatory variability in the imitation and retrieval tasks, children were separated into two groups based on performance on the BBTOP. Children with language impairment (based on performance on the SPELT-P 2/SPELT-3) were excluded from this analysis. The remaining children ($n = 29$) were

Figure 4. Relationship between language status and articulatory variability in imitation and retrieval. Symbols represent group means. Error bars represent standard error of the mean. TD = typically developing; LI = language impairment.

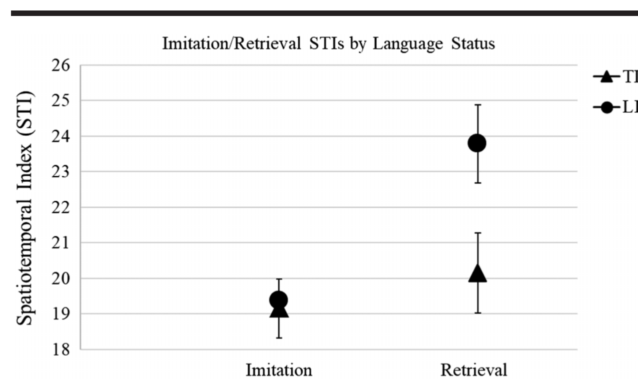


Table 3. Means and standard deviations by language status, speech status, and group for the kinematic measures in the imitation and retrieval tasks.

Kinematic measures	Imitation	Retrieval
STI by language status		
Typical language (n = 29)	19.15 (4.67)	20.15 (6.40)
Language impairment (n = 17)	19.37 (4.87)	23.79 (6.11)
STI by speech status		
Typical speech (n = 14)	17.77 (4.27)	19.70 (5.82)
Speech impairment (n = 15)	20.10 (4.80)	22.46 (6.51)
STI by group		
TD (n = 14)	18.31 (4.33)	19.12 (6.00)
SSD (n = 14)	20.40 (4.75)	21.36 (5.77)
DLD (n = 12)	18.81 (5.01)	23.14 (5.93)
CAS (n = 6)	19.56 (5.00)	23.62 (8.20)
Duration by group		
TD (n = 14)	0.84 (0.09)	0.84 (0.13)
SSD (n = 14)	0.85 (0.11)	0.90 (0.15)
DLD (n = 12)	0.89 (0.11)	0.93 (0.18)
CAS (n = 6)	0.88 (0.15)	0.87 (0.12)

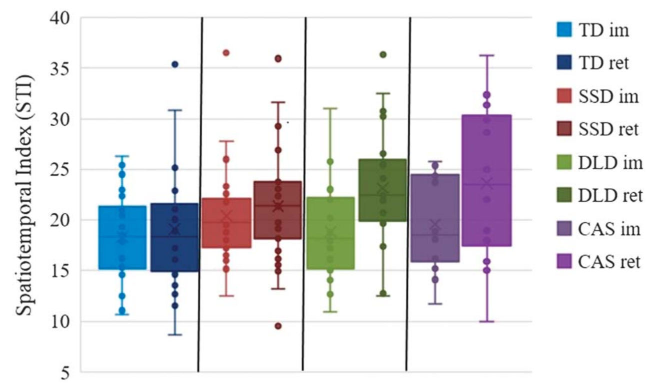
Note. STI = spatiotemporal index; TD = typically developing; SSD = speech sound disorder; DLD = developmental language disorder; CAS = childhood apraxia of speech.

coded as having either speech skills in the average range (n = 14; i.e., TD group) or speech impairment (n = 15; i.e., SSD group and one participant with CAS and typical language skills). Table 3 shows the STI means and standard deviations by speech status. A mixed ANOVA showed no main effect of speech status, $F(1, 27) = 1.80$, $p = .19$, $\eta_p^2 = .06$, no task effect, $F(1, 27) = 1.24$, $p = .28$, $\eta_p^2 = .04$, and no task by speech status interaction, $F(1, 27) = 0.03$, $p = .86$, $\eta_p^2 = .001$. Children with speech disorders and children with typical speech skills showed similar articulatory variability levels in imitated and retrieved phrases.

Articulatory Variability by Group: Exploratory Analysis of CAS

Finally, we conducted a relatively descriptive (due to small sample size) analysis to investigate whether children with CAS, because of a core deficit in planning or programming speech movement sequences, show high levels of articulatory variability. We predicted that, compared with children without motor speech impairments, children with CAS would show increased variability in the low-load imitation task. Similar to other published studies on CAS (e.g., Edeal & Gildersleeve-Neumann, 2011; Maas, Butalla, & Farinella, 2012; Maas & Farinella, 2012), our CAS group was quite small. Therefore, we were most interested in assessing group means and standard deviations and the effect sizes from this ANOVA (Edeal & Gildersleeve-Neumann, 2011; Gierut & Morrisette, 2011; Maas et al., 2012; Maas & Farinella, 2012). Table 3 shows the STI means and standard deviations by group, and Figure 5 shows the minimum, first quartile, median, third quartile, and maximum STI values by group in the imitation and retrieval tasks. A mixed ANOVA was used primarily to

Figure 5. Spatiotemporal index (STI) by group in imitation and retrieval. For each group, imitation is plotted on the left, and retrieval plotted on the right. TD = typically developing; SSD = speech sound disorder; DLD = developmental language disorder; CAS = childhood apraxia of speech.



obtain effect sizes. This test revealed no main effect of group, $F(3, 42) = 1.15$, $p = .34$, $\eta_p^2 = .08$. All children showed higher articulatory variability in the retrieval task, $F(1, 42) = 14.41$, $p < .001$, $\eta_p^2 = .26$, and there was no task by group interaction, $F(3, 42) = 1.86$, $p = .15$, $\eta_p^2 = .12$.

Interestingly, the task by group interaction effect size of 0.12 is similar to the interaction observed in the ANOVA by language status (effect size = 0.15), and the CAS and DLD groups produced similar mean STIs in the retrieval task. These results, though obtained from a small number of children, suggest that the CAS group patterns more closely with the DLD group. A clear motor impairment is not evident, at least in this small sample of children with CAS.

We also examined the durations of the extracted productions by group. Table 3 shows the duration means and standard deviations by group in the imitation and retrieval tasks. There was no main effect of group, $F(3, 42) = 0.74$, $p = .53$, $\eta_p^2 = .05$, no task effect, $F(1, 42) = 1.87$, $p = .18$, $\eta_p^2 = .04$, and no task by group interaction, $F(3, 42) = 1.09$, $p = .36$, $\eta_p^2 = .07$. Children from all four groups produced imitated and retrieved phrases with similar durations. Again, though the number of children with CAS was small, there was no evidence for a speech motor planning evidence in these six children.

Discussion

The goal of the current study was to investigate how language load (imitation vs. retrieval) affects articulatory variability as a function of language skill, speech skill, and diagnostic category. In the imitation task, somewhat surprisingly, all children showed similar articulatory variability levels, irrespective of language status, speech status, or group classification. When required to generate agent + action phrases in the retrieval task, all children showed higher articulatory variability. However, children with language impairment (i.e., DLD and CAS + LI) showed a disproportionate disruption in articulatory stability with increased retrieval load relative to peers with typical language skills.

All children produced more variable speech movements in the retrieval task. This finding replicates that of Vuolo and Goffman (2017) in a larger sample of children and contributes to a body of research documenting interactions between language and articulatory processing levels (e.g., Heisler et al., 2010; Rapp & Goldrick, 2006). As language load increases and processing demands become more complex (with lexical and phonological supports removed), articulatory control is perturbed. This result is consistent with other studies documenting that increased length (Sadagopan & Smith, 2008), syntactic complexity (Kleinow & Smith, 2000), articulatory complexity (Case & Grigos, 2016), and prosodic complexity (Goffman, 1999, 2004) disrupt articulatory control across a variety of talkers.

Speech kinematics and language load can also inform how language processing demands affect children with language impairment. Previous research has demonstrated that children with DLD show higher articulatory variability compared with TD peers (e.g., Brumbach & Goffman, 2014; Goffman, 1999), though it is not clear whether these motor speech difficulties are due to a weakness at higher language levels or reflect a comorbid weakness in speech motor planning or implementation. The results from the current study suggest that children with DLD do not show a low-level motor impairment—articulatory variability did not distinguish children with DLD from peers with typical language skills in the low-load imitation task. Rather, children with language impairment show high levels of articulatory variability in the retrieval task because of the increased language processing demands associated with active language generation. Language skill mediates the relationship observed between language load and speech motor control.

One caveat to the observed interaction between language status and language processing demands is that we ran this ANOVA by language status, which resulted in the inclusion of children with SSD (i.e., DLD + SSD) and CAS in this analysis. The overlap in diagnoses, particularly at young ages, makes it difficult to recruit "clean" diagnostic categories. Rather, a young child who presents with a deficit in a single domain is the exception rather than the rule (e.g., see Hill, 2001). We do acknowledge, however, that children with DLD who show concomitant speech difficulties may represent a distinct subgroup of children with DLD (though, in actuality, this subgroup represents the majority of clinically selected preschool-age children with DLD; see Alt et al., 2004; Deevy et al., 2010; Gray, 2006; Leonard, 2014; Vuolo et al., 2017).

Alternatively, if we consider these data from a slightly different perspective, as children with SSD or CAS who may or may not have co-occurring language deficits, it still is the presence of language deficits that drives the increased articulatory variability observed in a more challenging language generation task. From either the DLD or SSD perspective, it is the language deficit that drives the increases in articulatory variability and only in a relatively high-load language retrieval context.

There are several possible sources of difficulty that may account for the disproportionate disruption in articulatory control experienced by children with language impairment during the retrieval task. One possibility is that children with language impairment encode the language stimuli more shallowly during the learning/exposure phase (i.e., the imitation task in this study). Children with DLD are known to develop weak semantic representations of words relative to peers (Kail, Hale, Leonard, & Nippold, 1984; McGregor, Newman, Reilly, & Capone, 2002). At the lexical access stage, children with DLD also show word finding difficulties (Seiger-Gardner & Schwartz, 2008) and greater interference from competing words (Mainela-Arnold, Evans, & Coady, 2008). Even when accurately producing relatively simple sentences, such as "the mouse is eating the cheese," school-age children with DLD show subtle language formulation deficits compared with TD peers (Finneran, Leonard, & Miller, 2009). Language generation difficulties may propagate to speech motor planning and implementation stages, perturbing articulatory control.

These data are also consistent with the view that the declarative and procedural memory systems are involved in different aspects of language production. Rote imitation presumably relies more on declarative memory processes, which are a relative strength for children with language impairment. Indeed, imitation performance did not differentiate any of the diagnostic categories. The current study does, however, indicate that language retrieval and its associated syntactic, morphological, and phonological sequencing requirements are components of procedural learning that are affected in children with language impairment.

In contrast to DLD, children with CAS are thought to have a motor-based speech deficit. Therefore, it is reasonable to predict that children with CAS would produce phrases with high levels of articulatory variability even in a low-load imitation task. However, counter to our prediction, all children produced similar levels of articulatory variability in the imitation task. In the retrieval task, the similar effect size observed in the CAS group compared with the language impaired group suggests that language, not motor speech skill, drives the observed disruption in articulatory control associated with increased language load. Importantly, children with CAS also did not produce phrases with longer durations in either task. There are a few caveats to these findings that must be considered. An obvious limitation of the current study is that only a small number of children with CAS with heterogeneous apraxia features participated in our study, and this small n resulted in a lack of statistical power. In addition, though variability is an expected locus of deficit based on the presumed motor planning deficit associated with CAS, there may be other kinematic measures not included in this study that differentiate children with CAS from nonapraxic peers. Finally, the imitation task may not have been sufficiently difficult to capture group differences. The objective was to present very simple two-word phrases that even the most impaired talkers could produce. That goal was achieved, but the final language task was relatively low in complexity.

Children with SSD served as an important comparison group, as they show speech sound but not language deficits. Remarkably little work has investigated speech motor control in SSD, but it is reasonable to predict that these children would show motor implementation deficits compared with TD peers. However, children with SSD and typical language skills did not show higher levels of articulatory variability in either load condition compared with TD peers. The results of the current study suggest that speech sound skill does not directly correspond to articulatory movement patterning, at least in a simple phrase imitation or retrieval task.

Turning to intervention considerations, the retrieval task revealed clear differences between children with and without language difficulties. Disruptions in articulatory control may reflect a U-shaped learning trajectory (Gershkoff-Stowe & Thelen, 2004), in which articulatory precision is initially disrupted during new learning, but practice allows the motor system to reorganize and eventually restabilize. From this perspective, active language retrieval may facilitate language processing in children with language difficulties. It is possible that, with a single practice session, we were only able to observe the initial portion of the U-shaped learning curve, and if provided additional practice, articulatory variability would decrease in children with language impairment. A body of research has documented that retrieval practice, especially when the practice is sufficiently effortful, results in robust learning (see Roediger & Karpicke, 2006, for a review); this work should be extended to language and motor learning in disordered populations. Furthermore, children with DLD require more practice trials to reach performance levels comparable to those of TD peers in some procedural motor tasks (Lum, Conti-Ramsden, Morgan, & Ullman, 2014), which supports the notion of procedural memory impairments in DLD. Future work should continue to explore how active language retrieval facilitates learning and generalization in children with language impairment, such as those with DLD and CAS.

In contrast to retrieval practice, imitation practice may not be an appropriate intervention strategy for children with speech and/or language disorders. This is because imitation practice may not tax linguistic and motor speech processes to the extent necessary to effect change in intervention. In the current study, imitation did not differentiate children regardless of language skill, speech skill, or diagnostic category. In a study by Vuolo and Goffman (2017), imitation practice over five sessions actually resulted in increased articulatory variability levels from Session 1 to Session 5. Imitation may not promote learning that can be generalized across contexts.

In summary, these results contribute to a body of literature demonstrating that language (lexical and grammatical) and articulatory levels interact during speech production (e.g., Heisler et al., 2010; Rapp & Goldrick, 2006). Speech kinematics provide a sensitive index of the effects of language load and provide a promising avenue for future studies to investigate the relationship between retrieval practice and language skill. Future work should include a larger number of CAS participants and incorporate additional kinematic measures to determine the precise nature of the deficits in DLD and CAS and, perhaps, in SSD. Currently, in children with speech and/or language impairments, there is far more evidence that deficits in articulatory control are mediated by language rather than speech processes.

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders Grant F31 DC015176 (awarded to Vuolo) and Grant R01 DC04826 (awarded to Goffman). We thank Sara Benham, Janna Berlin, Lakin Brown, and Katie Isbell for their assistance with data collection and analyses. These data were collected at Purdue University.

References

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

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Technical Report]. Retrieved from http://www.asha.org/policy

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language, 61*(3), 438–456.

Bankson, N. W., & Bernthal, J. E. (1990). *Bankson–Bernthal Test of Phonology (BBTOP)*. Chicago, IL: Riverside Press.

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (Version 5.3.56) [Computer program]. Retrieved from http://www.praat.org/

Brumbach, A. C. D., & Goffman, L. (2014). Interaction of language processing and motor skill in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*(1), 158–171. https://doi.org/10.1044/1092-4388(2013/12-0215)

Case, J., & Grigos, M. I. (2016). Articulatory control in childhood apraxia of speech in a novel word–learning task. *Journal of Speech, Language, and Hearing Research, 59*(6), 1253–1268.

Dawson, J. I., Stout, C., Eyer, J. A., Tattersall, P., Fonkalsrud, J., & Croley, K. (2003). *Structured Photographic Expressive Language Test–Third Edition (SPELT-3)*. DeKalb, IL: Janelle Publications.

Dawson, J. I., Stout, C., Eyer, J. A., Tattersall, P., Fonkalsrud, J., & Croley, K. (2005). *Structured Photographic Expressive Language Test–Preschool 2 (SPELT-P 2)*. DeKalb, IL: Janelle Publications.

Deevy, P., Weil, L. W., Leonard, L. B., & Goffman, L. (2010). Extending use of the NRT to preschool-age children with and without specific language impairment. *Language, Speech, and Hearing Services in Schools, 41*(3), 277–288. https://doi.org/10.1044/0161-1461(2009/08-0096)

Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting

word repetition from picture naming. *Journal of Memory and Language, 56*(4), 490–520.
Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology (DEAP)*. San Antonio, TX: Psychological Corporation.
Edeal, D. M., & Gildersleeve-Neumann, C. E. (2011). The importance of production frequency in therapy for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 20*(2), 95–110.
Ehrler, D. J., & McGhee, R. L. (2008). *Primary Test of Nonverbal Intelligence (PTONI)*. Austin, TX: Pro-Ed.
Finneran, D. A., Leonard, L. B., & Miller, C. A. (2009). Speech disruptions in the sentence formulation of school-age children with specific language impairment. *International Journal of Language and Communication Disorders, 44*(3), 271–286.
Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49,* 87–115.
Gershkoff-Stowe, L., & Thelen, E. (2004). U-shaped changes in behavior: A dynamic systems perspective. *Journal of Cognition and Development, 5*(1), 11–36.
Gierut, J. A., & Morrisette, M. L. (2011). Effect size in clinical phonology. *Clinical Linguistics & Phonetics, 25*(11–12), 975–980.
Goffman, L. (1999). Prosodic influences on speech production in children with specific language impairment and speech deficits: Kinematic, acoustic, and transcription evidence. *Journal of Speech, Language, and Hearing Research, 42*(6), 1499–1517.
Goffman, L. (2004). Kinematic differentiation of prosodic categories in normal and disordered language development. *Journal of Speech, Language, and Hearing Research, 47*(5), 1088–1102. https://doi.org/10.1044/1092-4388(2004/081)
Goffman, L. (2010). Dynamic interaction of motor and language factors in development. In B. Maassen, P. H. H. M. Van Lieshout, R. Kent, & W. Hulstijn (Eds.), *Speech motor control: New developments in applied research* (pp. 137–152). England: Oxford University Press.
Goffman, L., & Smith, A. (1999). Development and phonetic differentiation of speech movement patterns. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 649–660.
Goldrick, M., Baker, H. R., Murphy, A., & Baese-Berk, M. (2011). Interaction and representational integration: Evidence from speech errors. *Cognition, 121*(1), 58–72.
Gray, S. (2006). The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 49*(5), 955–969.
Green, J. R., Moore, C. A., & Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research, 45*(1), 66–79. https://doi.org/10.1044/1092-4388(2002/005)
Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test–Preschool: Second Edition. *Language, Speech, and Hearing Services in Schools, 40*(2), 150–160.
Grigos, M. I., & Kolenda, N. (2010). The relationship between articulatory control and improved phonemic accuracy in childhood apraxia of speech: A longitudinal case study. *Clinical Linguistics & Phonetics, 24*(1), 17–40.
Grigos, M. I., Moss, A., & Lu, Y. (2015). Oral articulatory control in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 58*(4), 1103–1118. https://doi.org/10.1044/2015_JSLHR-S-13-0221
Haebig, E., Leonard, L., Usler, E., Deevy, P., & Weber, C. (2018). An initial investigation of the neural correlates of word processing in preschoolers with specific language impairment. *Journal of Speech, Language, and Hearing Research, 61*(3), 729–739. https://doi.org/10.1044/2017_JSLHR-L-17-0249
Heisler, L. & Goffman, L. (2016). The influence of phonotactic probability and neighborhood density on children's production of newly learned words. *Language Learning and Development, 12*(3), 338–356.
Heisler, L., Goffman, L. & Younger, B. (2010). Lexical and articulatory interactions in children's language production. *Developmental Science, 13*(5), 722–730.
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 and Communication Disorders, 36*(2), 149–171.
Hsu, H. J., & Bishop, D. V. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Developmental Science, 17*(3), 352–365.
Kail, R., Hale, C. A., Leonard, L. B., & Nippold, M. A. (1984). Lexical storage and retrieval in language-impaired children. *Applied Psycholinguistics, 5*(1), 37–49.
Kleinow, J., & Smith, A. (2000). Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 43*(2), 548–559. https://doi.org/10.1044/jslhr.4302.548
Kurland, J., Reber, A., & Stokes, P. (2014). Beyond picture naming: Norms and patient data for a verb-generation task. *American Journal of Speech-Language Pathology, 23*(2), S259–S270.
Lee, L. L., & Canter, S. M. (1971). Developmental sentence scoring: A clinical procedure for estimating syntactic development in children's spontaneous speech. *Journal of Speech and Hearing Disorders, 36*(3), 315–340.
Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–75.
Lewis, B. A., Freebairn, L. A., Hansen, A., Taylor, H. G., Iyengar, S., & Shriberg, L. D. (2004). Family pedigrees of children with suspected childhood apraxia of speech. *Journal of Communication Disorders, 37*(2), 157–175.
Lum, J. A., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex, 51,* 1–10.
Lum, J. A., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex, 48*(9), 1138–1154.
Maas, E., Butalla, C. E., & Farinella, K. A. (2012). Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 21*(3), 239–257.
Maas, E., & Farinella, K. A. (2012). Random versus blocked practice in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55*(2), 561–578. https://doi.org/10.1044/1092-4388(2011/11-0120)
Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2008). Lexical representations in children with SLI: Evidence from a frequency-manipulated gating task. *Journal of Speech, Language, and Hearing Research, 51*(2), 381–393.
Maner, K. J., Smith, A., & Grayson, L. (2000). Influences of utterance length and complexity on speech motor performance in children and adults. *Journal of Speech, Language, and Hearing Research, 43*(2), 560–573.
MathWorks. (2009). *MATLAB: High performance numeric computation and visualization software*. Natick, MA: Author.

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45*(5), 998–1014.

McPhillips, M., Finlay, J., Bejerot, S., & Hanley, M. (2014). Motor deficits in children with autism spectrum disorder: A cross-syndrome study. *Autism Research, 7*(6), 664–676.

Moss, A., & Grigos, M. I. (2012). Interarticulatory coordination of the lips and jaw in childhood apraxia of speech. *Journal of Medical Speech-Language Pathology, 20*(4), 127–132.

Nip, I. S., Green, J. R., & Marx, D. B. (2011). The co-emergence of cognition, language, and speech motor control in early development: A longitudinal correlation study. *Journal of Communication Disorders, 44*(2), 149–160.

Nittrouer, S. (1993). The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study. *Journal of Speech and Hearing Research, 36,* 959–972.

Rapp, B., & Goldrick, M. (2006). Speaking words: Contributions of cognitive neuropsychological research. *Cognitive Neuropsychology, 23*(1), 39–73.

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52,* 271–277.

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181–210.

Rvachew, S., & Brosseau-Lapré, F. (2016). *Developmental phonological disorders: Foundations of clinical practice* (2nd ed.). San Diego, CA: Plural.

Sadagopan, N., & Smith, A. (2008). Developmental changes in the effects of utterance length and complexity on speech movement variability. *Journal of Speech, Language, and Hearing Research, 51*(5), 1138–1151.

Saletta, M., Goffman, L., Ward, C., & Oleson, J. (2018). Influence of language load on speech motor skill in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 61*(3), 675–689.

Sanjeevan, T., & Mainela-Arnold, E. (2017). Procedural motor learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 60*(11), 3259–3269.

Schopler, E., Reichler, R. J., & Renner, B. R. (2010). *The Childhood Autism Rating Scale–Second Edition (CARS-2)*. Los Angeles, CA: Western Psychological Services.

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

Shriberg, L. D., Potter, N. L., & Strand, E. A. (2011). Prevalence and phenotype of childhood apraxia of speech in youth with galactosemia. *Journal of Speech, Language, and Hearing Research, 54*(2), 487–519.

Smith, A., & Goffman, L. (1998). Stability and patterning of speech movement sequences in children and adults. *Journal of Speech, Language, and Hearing Research, 41*(1), 18–30.

Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104,* 493–501.

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology, 45*(1), 22–33.

Strand, E. A., Stoeckel, R., & Baas, B. (2006). Treatment of severe childhood apraxia of speech: A treatment efficacy study. *Journal of Medical Speech-Language Pathology, 14*(4), 297–307.

Terband, H., Maassen, B., Van Lieshout, P. H. H. M., & Nijland, L. (2011). Stability and composition of functional synergies for speech movements in children with developmental speech disorders. *Journal of Communication Disorders, 44*(1), 59–74.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41*(3), 399–433.

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews, 51,* 205–222.

van der Merwe, A. (2011). A speech motor learning approach to treating apraxia of speech: Rationale and effects of intervention with an adult with acquired apraxia of speech. *Aphasiology, 25,* 1174–1206.

Vuolo, J., & Goffman, L. (2017). An exploratory study of the influence of load and practice on segmental and articulatory variability in children with speech sound disorders. *Clinical Linguistics & Phonetics, 31*(5), 331–350.

Vuolo, J., Goffman, L., & Zelaznik, H. N. (2017). Deficits in coordinative bimanual timing precision in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 60*(2), 393–405.

Walsh, B., & Smith, A. (2002). Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research, 45*(6), 1119–1133.

Zelaznik, H. N., & Goffman, L. (2010). Generalized motor abilities and timing behavior in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(2), 383–393.