**Research Article**

# Lexical and Phonological Variability in Preschool Children With Speech Sound Disorder

Toby Macrae,[a] Ann A. Tyler,[b] and Kerry E. Lewis[c]

**Purpose:** The authors of this study examined relationships between measures of word and speech error variability and between these and other speech and language measures in preschool children with speech sound disorder (SSD).
**Method:** In this correlational study, 18 preschool children with SSD, age-appropriate receptive vocabulary, and normal oral motor functioning and hearing were assessed across 2 sessions. Experimental measures included word and speech error variability, receptive vocabulary, nonword repetition (NWR), and expressive language. Pearson product–moment correlation coefficients were calculated among the experimental measures.
**Results:** The correlation between word and speech error variability was slight and nonsignificant. The correlation between word variability and receptive vocabulary was moderate and negative, although nonsignificant. High word

variability was associated with small receptive vocabularies. The correlations between speech error variability and NWR and between speech error variability and the mean length of children's utterances were moderate and negative, although both were nonsignificant. High speech error variability was associated with poor NWR and language scores.
**Conclusion:** High word variability may reflect unstable lexical representations, whereas high speech error variability may reflect indistinct phonological representations. Preschool children with SSD who show abnormally high levels of different types of speech variability may require slightly different approaches to intervention.

**Key Words:** speech sound disorder, speech variability, word variability, error variability, inconsistency

Historically, speech variability has received much less attention than speech accuracy in studies of typical and disordered speech development in young children. Studying speech variability in addition to accuracy is clinically important for several reasons. First, although a relationship clearly exists between speech sound accuracy and speech intelligibility, the correlation is modest ($r = .42$; Shriberg & Kwiatkowski, 1982). Shriberg and Kwiatkowski (1982) concluded that "speech intelligibility reflects a complex of factors in addition to articulation proficiency" (p. 264). In adult speakers with dysarthria, increased variability in the production of vowels and syllables is associated with decreased speech intelligibility and increased severity of impairment (Kim, Hasegawa-Johnson, & Perlman, 2010; Ziegler, Hartmann, & Hoole, 1993). Young children

with speech sound disorders (SSD) become more intelligible as listeners become familiar with their speech error patterns (Shriberg & Kwiatkowski, 1982). Children with variable speech are likely to be less intelligible than children with consistent speech because of the unpredictability in their speech productions (Holm, Crosbie, & Dodd, 2005).

Speech variability is also clinically important because word variability (one type of speech variability) is a core feature of certain subtypes of SSD in children—namely, childhood apraxia of speech (CAS; American Speech-Language-Hearing Association [ASHA], 2007) and inconsistent disorder (Dodd, 2005). *Word variability* refers to variability in repeated productions of the same word. For example, a child may produce the target word *cat* as [kæt], [dæt], and [dæ] on three different occasions in conversation during a play session. Children with these subtypes of SSD may require unique approaches to intervention. For example, children with CAS may benefit from an approach based on motor learning theory, such as the Nuffield Centre Dyspraxia Programme (3rd ed.; see Williams & Stephens, 2010). Children with inconsistent disorder may benefit from an approach that focuses on consistent production of a core set of vocabulary items (e.g., core vocabulary intervention; Dodd, Holm, Crosbie, & McIntosh, 2010).

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

The third reason for the clinical importance of speech variability is that it may relate to prognosis in treatment for children with SSD. Forrest and colleagues (Forrest, Dinnsen, & Elbert, 1997; Forrest, Elbert, & Dinnsen, 2000) studied the effect of speech error variability (another type of speech variability) on speech sound learning and generalization in children with SSD. *Speech error variability* refers to variability in speech sounds that are produced in place of misarticulated targets. Targets with a consistent speech error are those that are misarticulated and for which only one speech sound is produced in their place. For example, a child may produce the target sound /tʃ/ as [ts] in *chair* ([tsɛr]), *church* ([tsɝˑts]), and *patch* ([pæts]). Targets with variable speech errors are those that are misarticulated and for which several speech sounds are produced in their place. For example, a child may produce the target /s/ as [t] in *saw* ([tɔ]), as [d] in *sun* ([dʌn]), and as an omission in *bus* ([bʌ]). Children receiving treatment for a target produced with a consistent error showed greater learning and generalization of the target than children receiving treatment for a target produced with variable errors (Forrest et al., 1997, 2000).

Studying speech variability is also theoretically important. Dynamic systems theory is a theory of action that was formulated to account for real-time changes in motor behavior in infants (Thelen & Bates, 2003). This theory was extended to account for longer term changes in motor development. A dynamic systems view of development "considers the origins and functions of variability as absolutely central for understanding change" (Thelen & Smith, 1994, p. 145). Variability is associated with transitions between developmental stages and is viewed as a "potential driving force of development and a potential indicator of ongoing processes" (van Geert & van Dijk, 2002, p. 341). When lying on their backs, for example, newborn infants perform highly coordinated alternating leg kicks. At about 1 month of age, coordination between the legs becomes highly variable. This variability leads to new forms of coordination between the legs, for example, simultaneous kicking (Thelen & Smith, 1994).

Thelen and Smith (1994) suggested that infants must free themselves of the stable patterns of the newborn period before they can assemble new patterns of coordination. The variability present during developmental transitions provides infants with a wide array of coordinative possibilities. In other words, variability in motor development allows infants to explore new patterns of motor behavior. Variability is characteristic of the development of other biological and psychological systems, including speech-language development (van Geert & van Dijk, 2002). Dynamic systems theory has potential in explaining the variability that is seen in speech-language development. Peaks in variability in speech production may presage developmental change, just as variability has been shown to precede real-time changes in motor behaviors and longer term changes in other areas of development.

## Word Variability

Sosa and Stoel-Gammon (2006) studied whole-word variability in four children between the ages of 1 and 2 years. They used Ingram's (2002) measure of variability, the proportion of whole-word variation, which is calculated by dividing the number of different phonetic forms of a word by the total number of productions of that word in a speech sample. The authors found that variability fluctuated throughout the duration of the study and peaked when children had acquired ~150–200 words and when two-word combinations were first observed. Sosa and Stoel-Gammon suggested that this peak in variability reflected a reorganization of the linguistic system that included a transition from underlying holistic to phonemic representations as well as the analysis and combination of individual words. Underlying representations are "part of the mental lexicon that stores the information needed to recognize and produce words" (Stoel-Gammon, 2011, p. 17).

Sosa and Stoel-Gammon's (2006) interpretation is consistent with Metsala and Walley's (1998) lexical restructuring model (LRM). According to this model, infants' early lexical representations are holistic in nature as there is simply no need to represent words in a more detailed manner. As children's vocabularies grow, however, the increasing similarity among words in the lexicon creates pressure to form more fine-grained, phonemic representations to allow for accurate word recognition and production, and this continues into middle childhood (Metsala & Walley, 1998).

Metsala and Walley (1998) summarized findings from the speech perception literature to support their model. First, phonemic perception shows a protracted course of development in children and may continue into young adulthood. Children's ability to identify and discriminate phonemes improves as their vocabularies grow, and this is thought to reflect more clearly defined phonemic categories. Second, adults are more sensitive than children to the phonemic composition of words in word recognition tasks and are more adept at identifying a spoken word when only part of the word is presented to them. Children need to perceive the entire word if they are to access its underlying representation.

Although word variability has been shown to peak during developmental change, studies point to a general trend of decreasing word variability throughout both typical and disordered speech development (Burt, Holm, & Dodd, 1999; Holm, Crosbie, & Dodd, 2007; Iuzzini & Forrest, 2011). Burt et al. (1999) found a significant negative correlation between age and word variability in children with typical speech development, ages 3;10 (years;months) to 4;10. These authors measured variability in the production of the same words across different linguistic contexts (e.g., imitation, spontaneous production in response to picture prompts, and spontaneous production in connected speech). Holm et al. (2007) observed a significant decrease in word variability with increasing age in children with typical speech development, ages 3;0 to 6;11. These authors measured variability using Dodd's (1995) Inconsistency Assessment. This assessment requires children to name 25 colored pictures on three separate occasions within a session. The youngest group, ages 3;0–3;5, showed significantly more variability in their word productions than all other groups. An average of 13% of target words was produced variably by children in the

youngest group. Iuzzini and Forrest (2011) found decreasing word variability with increasing age in children with typical and disordered speech development, ages 3;0 to 5;8. In this study, variability reflected the proportion of target words that were produced variably across three or more productions, although it is unclear what linguistic context was used.

Within an overall trend of decreasing word variability in children with typical speech development, peaks may presage developmental change. In children with SSD, however, persistent variability, or variability that is higher than for other children with SSD, may be characteristic of a subgroup with a unique underlying deficit. In children with SSD, word variability is most often associated with CAS. CAS is characterized by three key features, one of which is variability. Specifically, CAS is associated with "(a) inconsistent errors on consonants and vowels in repeated productions of syllables or words, (b) lengthened and disrupted coarticulatory transitions between sounds and syllables, and (c) inappropriate prosody, especially in the realization of lexical or phrasal stress" (ASHA, 2007, p. 2). CAS is thought to reflect an impairment in speech output, specifically, "planning and/or programming spatiotemporal parameters of movement sequences" (ASHA, 2007, p. 1). This motor planning and/or programming deficit may be responsible for the high levels of word variability that are seen in children with CAS (Marquardt, Jacks, & Davis, 2004).

Dodd (2005) proposed a classification system of children with SSD that includes subgroups characterized according to word variability. According to Dodd, children who use developmental and nondevelopmental phonological processes and show variable productions of less than 10 of the 25 target words on Dodd's (1995) Inconsistency Assessment are identified as having *consistent disorder,* and children who show variable productions of 10 or more of the 25 target words are identified as having *inconsistent disorder.* Whereas a diagnosis of CAS is based on other features in addition to word variability, a diagnosis of inconsistent disorder is based entirely on word variability.

Dodd and colleagues (Bradford & Dodd, 1994; Holm, Farrier, & Dodd, 2008) conducted several studies in an attempt to identify unique deficits in children with inconsistent disorder. Children in this subgroup showed age-appropriate oral motor functioning but deficits in spontaneously producing, imitating, and spelling nonwords and real words. In the most recent study, Holm et al. (2008) pointed to a "linguistic breakdown rather than a motor planning and implementation disorder" (p. 312). Specifically, the authors proposed a deficit in selecting and assembling phonemes to form words, which they referred to as *phonological assembly,* to account for the deficits and highly variable word productions that are seen in children with inconsistent disorder. More research is required to determine if high word variability in children with SSD other than CAS reflects an underlying deficit and, if so, what this deficit might be.

Relationships between measures of word variability and other measures of speech and language abilities may point to underlying deficits. Several studies that have examined the relationship between word variability and vocabulary knowledge have yielded mixed results. Sosa and Stoel-Gammon (2006) found a negligible relationship between word variability and expressive vocabulary in children with typical speech development. Holm et al. (2008) found that children in their inconsistent disorder subgroup had receptive vocabulary scores that were no different than those of children in the other subgroups. More recently, however, Sosa and Stoel-Gammon (2012) found a high negative correlation ($r = -.81$) between word variability and expressive vocabulary. The conflicting findings across Sosa and Stoel-Gammon's two studies may be due to their different measures of word variability. Sosa and Stoel-Gammon (2006) calculated overall word variability as a ratio of the number of different phonetic forms of all words produced two or more times to the total number of productions of these words. This measure does not account for participants attempting different numbers of target words. For example, a child might produce three target words four times each, with the exact same phonetic form for each production of each word. This yields a word variability ratio of 3:12, which equals .25. Another child might produce six target words two times each, also with the exact same phonetic form for each production of each word. This yields a word variability ratio of 6:12, which equals .50. Both of these children are completely consistent in their word productions, yet they receive quite different variability scores. Sosa and Stoel-Gammon (2012) controlled for different numbers of target words across participants by calculating the mean number of different phonetic forms per target word. This measure, therefore, would seem to be a more valid reflection of word variability than the measure that was employed in the earlier study. The relationship between word variability and vocabulary knowledge requires further exploration. Negative correlations may point to unstable lexical representations in children with high word variability.

### Speech Error Variability

Studies examining the relationship between speech error variability and phonological change in children with SSD provide some insight into the nature of the deficit underlying this type of variability. Recall that targets with a consistent speech error have shown greater gains following treatment than targets with variable speech errors (Forrest et al., 1997, 2000). Variable speech errors may reflect a lack of categorical representation for a target sound, and "the development of categories is a prerequisite to phonological learning" (Forrest et al., 1997, p. 74). A later study, however, revealed that speech error variability calculated for all sounds produced in error predicted phonological change such that children with the most variable substitutes showed the greatest gains (Tyler, Lewis, & Welch, 2003). This relationship is in the opposite direction to what would be expected based on Forrest and colleagues' studies and may be due to methodological differences across the studies. First, Tyler et al. (2003) included children with SSD and expressive language impairments, whereas only some of Forrest and

colleagues' participants had documented SSD, and all had age-appropriate expressive language abilities. Second, Tyler et al. measured speech error variability and phonological change across the entire phonological system. Speech error variability was measured using the Error Consistency Index (ECI; Tyler et al., 2003), which reflects the total number of different speech errors for all consonant targets in a single-word elicitation task. In contrast, Forrest and colleagues measured speech error variability and phonological change for individual targeted sounds. Third, participants in the Forrest et al. studies received treatment for one singleton fricative that was omitted from the phonetic inventory and for which participants produced variable speech errors. In contrast, participants in the Tyler et al. study received treatment for anywhere between four and eight singleton or cluster targets. When selecting treatment targets, these authors did not control for the variability of speech errors. These methodological differences make it difficult to draw conclusions about the effect of speech error variability on phonological change in children with SSD and are likely responsible for the seemingly conflicting findings across the studies. More research is needed to study speech error variability in children with SSD.

Significant correlational relationships may elucidate the deficits underlying high speech error variability in children with SSD. Along these lines, Preston and Koenig (2011) examined relationships among measures of word, speech error, and phonetic variability in school-age children with residual speech errors. Speech error variability was highly correlated with word variability and moderately correlated with two measures of phonetic variability (i.e., word duration variability and voice onset time variability). Although variability in vowel formants, another measure of phonetic variability, was moderately correlated with word and speech error variability, the correlations were negative. That is, children with high variability in their vowel formants showed low word and speech error variability. Low or negligible correlations were observed among acoustic measures of phonetic variability and among all other measures of variability. One explanation given for the lack of convergence among transcription-based and acoustic measures of variability was that transcription-based measures may reflect phonological-level processing, whereas acoustic measures may reflect speech motor functioning. Iuzzini and Forrest (2011) examined relationships among measures of word and speech error variability in children with and without SSD, including those with CAS. Like Preston and Koenig, Iuzzini and Forrest found high correlations between measures of word and speech error variability.

Although nonword repetition (NWR) has been identified as a potential clinical marker of specific language impairment (SLI; Bishop, North, & Donlan, 1996; Dollaghan & Campbell, 1998), it may also provide insight into the deficits underlying high speech error variability. According to Munson, Kurtz, and Windsor (2005), NWR is a complex task involving several cognitive processes, including "perceiving and discriminating the acoustic signal, matching the signal with phonological representations in memory,

planning the articulatory movements required to replicate the nonword, and executing the response" (p. 1,033). With regard to phonological representations, Gathercole (2006) suggested that NWR "is influenced by the quality and persistence of the phonological representations that are characteristic of an individual…and by prior factors affecting the initial construction of the phonological representation" (p. 519). In order for speakers to be able to combine phonemes into unfamiliar strings during NWR tasks, therefore, they must possess, among other things, distinct phonological representations. Elbro, Borstrom, and Petersen (1998) referred to distinctness as "the relative distance between a phonological representation and its neighbors" (p. 40) and suggested that indistinct representations may be more easily confused with their neighbors than distinct representations. Rispens and Baker (2012) found that the distinctness of children's phonological representations was highly correlated with NWR in 5-year-old children with typical language development. In their study, both phonological short-term memory and phonological representations predicted significant amounts of variance in NWR in 5- to 8-year-olds with typical language, with phonological representations predicting a larger amount of variance (31%) than phonological short-term memory (10%). Sosa and Stoel-Gammon (2012) pointed out that no study has examined speech variability and NWR in the same group of children. Research is required to determine the relationship between speech error variability and NWR in children with SSD. Negative correlations may point to indistinct phonological representations in children with high speech error variability.

## Purpose

The purpose of the present study was to determine if word variability and phonological-level variability reflect unique underlying processes and, if so, what these processes might be. First, we examined the relationship between word and speech error variability. According to the LRM, the development of distinct phonological representations occurs only after children have acquired sufficient lexical representations. Based on the proposal that high word variability reflects unstable lexical representations, and high phonological variability reflects indistinct phonological representations, peaks in phonological variability would be expected to follow peaks in word variability. Therefore, we hypothesized that word and speech error variability would not be correlated. Second, we examined relationships between measures of speech variability and other speech and language measures—namely, receptive vocabulary, the standard score from a standardized test of expressive and receptive language, mean length of utterance (MLU), and NWR—to provide insight into what each measure of speech variability reflects. Although some speech variability may reflect developmental change, high levels of variability in children with SSD may be indicative of an underlying deficit. Negative correlations would support this. We hypothesized that word variability would be negatively correlated with receptive vocabulary, and speech error variability would be negatively correlated with NWR performance.

## Method

### Participants

Eighteen children (13 boys and 5 girls), ages 3;6–5;5 ($M_{age}$ = 4;8), participated in the study. All participants met the following inclusion criteria: (a) presence of SSD, as confirmed by a score at least 1 *SD* below the mean on the Bankson–Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990); (b) no existing diagnosis of CAS; (c) age-appropriate receptive vocabulary, as confirmed by a score > 1 *SD* below the mean on the Peabody Picture Vocabulary Test—III (PPVT–III; Dunn, Dunn, & Williams, 1997); (d) normal oral motor functioning, as confirmed by a score > 1 *SD* below the age-appropriate mean on at least one subtest of the Oral and Speech Motor Control Protocol (Robbins & Klee, 1987); and (e) normal hearing, as confirmed by positive responses to 1000-Hz, 2000-Hz, and 4000-Hz stimuli presented at 25 dB HL in audiometric testing.

Participants' descriptive statistics for the inclusion criteria are presented in Table 1. Potential participants were recruited from early childhood programs in the Washoe County School District, Reno, Nevada, and from the speech and hearing clinic at the University of Nevada, Reno (UNR). Letters of invitation were given to parents of children who were receiving treatment for SSD or waiting to receive an assessment for suspected SSD.

### Experimental Measures

All participants received a 2-hr assessment in order to determine if they met the inclusion criteria and to obtain results for the experimental measures. These included severity of SSD, as measured by percentage of consonants correct (PCC), word and speech error variability, NWR, and language abilities. All testing was conducted by the first author in the UNR speech and hearing clinic in a quiet room, across two sessions. Sessions were audio-recorded using a Sony ICD-P320 digital audio recorder at a sampling rate of 8 kHz.

*Word variability.* Word variability was calculated using a story retell. *Pancakes for Breakfast* (dePaola, 1978) is a wordless book with repetitive themes. A story script based on the book and designed for the purposes of this study was first read to each participant in its entirety at the very start of the assessment session. The story was again read to each participant, in six sections corresponding to six broad events in the story, toward the end of the session. At the end of each section, the examiner instructed the participant to retell that part of the story. The story script has 67 content words that are repeated at least once by the examiner. Content words produced three or more times by the participant were transcribed using International Phonetic Alphabet (IPA) broad transcription and were included in the analysis. All function words were excluded. Additionally, only words with CVC or more complex syllable structures were included. All words with more basic structures (i.e., V, CV, and VC) were excluded.

Word variability was calculated from the story retell using target variability (TV). TV reflects the proportion of target words with variable productions. It was calculated by dividing the number of content words produced three or more times with variable productions by the total number of content words produced three or more times. TV is the inverse of Marquardt et al.'s (2004) target stability measure, which reflects the proportion of target words for which all tokens are produced alike. Only consonant productions were considered when determining whether phonetic forms of a word were the same or different.

*Speech error variability.* Speech error variability was calculated using the ECI. The ECI was based on the responses from the BBTOP and 20 additional words that were selected to ensure that each of the 23 consonants was sampled three times each in initial and final word positions. Participants' responses were transcribed using IPA broad transcription. For each target consonant, the total number of different substitutions/omissions, regardless of word position, was summed. For example, /f/ occurred word initially in the words *fish, fire,* and *feather,* and word finally in the words *knife, leaf,* and *chief.* If a child produced [pɪʃ] for *fish,* [faɪr] for *fire,* [wɛdə] for feather, [naɪs] for knife, [lip] for leaf, and [ʧif] for chief, the number of different sound substitutions/omissions for /f/ across initial and final word positions was three. This process was repeated for each of the other 23 consonants. The number of different speech errors was summed across all 23 consonants to yield the ECI for each participant.

*NWR.* NWR was measured using the Syllable Repetition Task (SRT; Shriberg et al., 2009), which was designed for use with children with SSD. The SRT consists of 18 multisyllabic nonwords containing CV sequences: eight 2-syllable nonwords (CVCV), six 3-syllable nonwords (CVCVCV), and four 4-syllable nonwords (CVCVCVCV). The nonwords contain only early developing consonants (/b, d, m, n/) and one vowel (/ɑ/). As per Shriberg et al. (2009), scoring for the SRT was based on the PCC.

*Language.* Three measures of language were obtained. The first was the Core Language score from the Clinical Evaluation of Language Fundamentals—Preschool, Second Edition (CELF–P2; Semel, Wiig, & Secord, 2004). This is a standardized summary score based on participants'

**Table 1.** Participant summary data for inclusion criteria.

| Measure | *M* | *SD* | Range |
|---|---|---|---|
| BBTOP standard score | 67.89 | 3.68 | 65.00–76.00 |
| PPVT–III standard score | 106.72 | 10.65 | 86.00–126.00 |
| OSMCP TSS | 23.78 | 0.55 | 22.00–24.00 |
| OSMCP TFS | 106.50 | 3.43 | 22.00–24.00 |
| OSMCP PRR | 1.51 | 0.38 | 1.00–2.33 |

*Note.* BBTOP = Bankson–Bernthal Test of Phonology (Bankson & Bernthal; 1990); PPVT–III = Peabody Picture Vocabulary Test—III (Dunn et al., 1997); OSMCP = Oral and Speech Motor Control Protocol (Robbins & Klee, 1987); TSS = total structural score; TFS = total functional score; PRR = polysyllabic repetition rate.

performance on the Sentence Structure, Word Structure, and Expressive Vocabulary subtests of the CELF–P2. The second measure, MLU$_m$, reflects the average length in morphemes of participants' spoken utterances from the story retell. Participants' retell samples were transcribed and MLU$_m$ was calculated using the Systematic Analysis of Language Transcripts program (SALT; Miller & Chapman, 2000). The third measure of language was the standard score from the PPVT—III, the test of receptive vocabulary. Descriptive statistics for the experimental measures are provided in Table 2.

### Reliability

Inter- and intrarater reliability were assessed for the following: (a) identification of the participants' consonant productions in single words using IPA broad transcription, the basis of the majority of the experimental tasks, including the story retell, the BBTOP, the 20 additional words, and the SRT; and (b) transcription of the participants' spoken utterances from the story retell and subsequent calculation of MLU$_m$. For IPA transcription, audio recordings of BBTOP samples from two participants (11%) were selected at random. Interrater reliability was calculated as the percent agreement between the first author, who was the original transcriber, and a trained research assistant, who was a graduate student in communication science and disorders, in the identification of consonants in the BBTOP responses. Interrater reliability was 87.25%. Intrarater reliability was calculated as the percent agreement between the first author's original transcription and the repeated transcription. Intrarater reliability was 95.05%.

For the calculation of MLU$_m$, audio recordings of the story retells from two participants, selected at random, were retranscribed and the MLU$_m$ was recalculated using the SALT program. Interrater reliability was calculated as the mean absolute difference between the first research assistant's original MLU$_m$ and the second research assistant's recalculated MLU$_m$ across the two participants. The mean absolute difference was 0.22, which represents 4.9% of the original mean MLU$_m$ that was calculated across the two participants. Intrarater reliability was calculated as the mean absolute difference between the first research assistant's original and recalculated MLU$_m$ across the two participants. The mean absolute difference was 0.51, which represents 11.3% of the original mean MLU$_m$ that was calculated across the two participants.

## Results

The data in Table 2 reveal some important participant characteristics. First, the participants presented with SSD ranging in severity from mild–moderate to severe, with a mean PCC representing a moderate–severe rating (Shriberg & Kwiatkowski, 1982). Although PCC scores were fairly evenly distributed among mild–moderate ($n = 5$), moderate-severe ($n = 7$), and severe ($n = 6$) ratings, no participants

**Table 2.** Participant summary data for the experimental measures.

| Measure | M | SD | Range |
|---|---|---|---|
| PCC | 56.63 | 12.99 | 35.4–78.3 |
| TV | .41 | .20 | .15–.79 |
| ECI | 21.94 | 7.97 | 12–38 |
| SRT | 74.11 | 18.13 | 28–96 |
| CELF–P2 | 95.61 | 14.25 | 65–119 |
| MLU$_m$ | 5.00 | 1.69 | 2.10–7.74 |

*Note.* PCC = percentage of consonants correct; TV = target variability; ECI = Error Consistency Index (Tyler et al., 2003); SRT = PCC on the Syllable Repetition Task (Shriberg et al., 2009); CELF–P2 = core Language score from the Clinical Evaluation of Language Fundamentals—Preschool, Second Edition (Semel et al., 2004); MLU$_m$ = mean length of utterance in morphemes.

presented with mild SSD. Second, the majority of participants presented with age-appropriate language abilities. All participants scored > 1 *SD* below the mean on the PPVT–III (see Table 1); most participants (14 of 18 participants, 78%) also scored above this cutoff on the CELF–P2. Third, participants' TV scores ranged from .15 to .79, with a mean of .41. That is, 15% of the least variable child's target words and 79% of the most variable child's target words were produced with variable productions. This represents a wide range of word variability across the participants. The number of target words included in the word variability analysis ranged from 9 to 26 ($M = 20$), and the total number of word productions across all targets ranged from 43 to 146 ($M = 97$), across the participants.

Pearson product–moment correlation coefficients were calculated between the variability measures and between these and the other speech and language measures. The significance level was set at $p < .006$, based on a Bonferroni correction for multiple (nine) comparisons. Three of the six experimental measures rely, to some degree, on the accuracy of speech sound production, including the measures of word and speech error variability (i.e., TV and ECI). For two words or two sounds to be considered variable, at least one of the productions must be in error. Furthermore, the SRT is scored according to the PCC. In an effort to control for the effect of severity of SSD on these measures, partial correlations controlling for PCC were calculated. The correlation between TV and the ECI was slight (Guilford, 1956) and nonsignificant. The correlation between TV and the PPVT–III standard score was moderate, negative, and significant ($r = -.45$, $p = .034$). High word variability was associated with small receptive vocabularies. The correlations between the ECI and SRT ($r = -.44$, $p = .040$) and between the ECI and MLU$_m$ ($r = -.43$, $p = .042$) were moderate, negative, and significant. The correlation between the ECI and the Core Language score from the CELF–P2 was negative and approached moderate in size and significance ($r = -.39$, $p = .063$). All correlations were nonsignificant. High error variability was associated with poor SRT performance and low language scores.

# Discussion

## Relationship Between Word and Speech Error Variability

The first aim of this study was to examine the relationship between word and speech error variability in young children with SSD. The LRM proposes that infants' burgeoning lexicons precipitate increases in the level of detail they store in their underlying lexical representations. According to this model, the development of distinct phonological representations follows the acquisition of a sufficient number of lexical representations. Therefore, one would expect peaks in phonological-level variability to follow peaks in word variability. Although this study measured variability at only one point in time, the slight nonsignificant correlation between word and speech error variability is in line with the assumption that these two different types of speech variability reflect unique underlying processes and would show different developmental progressions, thereby supporting the LRM.

The slight correlation is in contrast to Preston and Koenig's (2011) and Iuzzini and Forrest's (2011) findings of high correlations between word variability and speech error variability. These differences may be due to the different populations of interest: preschoolers with SSD in the present study; school-age children with residual speech errors in Preston and Koenig, whose variability may have been more closely related to speech motor abilities than linguistic abilities; and preschoolers with and without SSD, including those with CAS, in Iuzzini and Forrest. In the children with CAS, at least, their high word variability was also likely related to their speech motor abilities. Preston and Koenig also suggested that different measures of variability may reflect unique underlying processes: "Transcriptional measures may capture a phonological level of representation/processing, whereas at least some of the acoustic measures…may better reflect phonetic or motoric processes" (p. 181). Preston and Koenig's transcriptional measures included word and phonological variability. It is suggested that word variability reflects lexical-level processing, whereas speech error variability reflects phonological-level processing. Specifically, high word variability reflects unstable lexical representations, whereas high speech error variability reflects indistinct phonological representations. The findings from the present study that support these proposals will now be discussed.

## Relationships Between Variability and Speech and Language Measures

The second aim of this study was to examine the relationships between measures of speech variability and other speech and language measures. Word variability was negatively correlated with receptive vocabulary, indicating that children with high word variability had small receptive vocabularies. This is consistent with Sosa and Stoel-Gammon's (2012) finding of a negative correlation and supports Holm et al.'s (2008) suggestion that children with high word variability have a linguistic deficit. These findings contribute to our understanding of the role of word variability in speech development. Sosa and Stoel-Gammon (2006) found that, in children with typical development, peaks in word variability presaged developmental change, which is in line with dynamic systems theory. The results of the present study suggest that, in preschoolers with SSD, high levels of word variability reflect an underlying deficit characterized by unstable lexical representations, which may be due to small vocabularies. Unstable lexical representations may be responsible for Holm et al.'s proposed deficit in phonological assembly in children with high word variability. Presumably, the more children use words and the more their vocabularies grow, the more stable their underlying representations will become.

In our study, speech error variability was negatively correlated with SRT; that is, children with the most variable speech errors showed the poorest SRT performance. This supports the proposal that variable speech errors, like poor SRT performance, are an overt manifestation of indistinct phonological representations (Forrest et al., 1997, 2000; Sosa & Stoel-Gammon, 2012; Tyler & Lewis, 2005). Sosa and Stoel-Gammon (2012) proposed that production variability and SRT may both assess "the degree of abstract phonemic knowledge" (p. 605). Forrest et al. (1997) suggested that children with a consistent error for a target sound may lack the ability to produce the sound, but they understand that the sound must be produced consistently in different contexts and word positions. That is, they understand the categorical nature of the sound. Categorical representation of speech sounds develops with increasing age (Coady, Evans, Mainela-Arnold, & Kluender, 2007; Hazan & Barrett, 2000; Liker & Gibbon, 2008; Mayo, Scobbie, Hewlett, & Waters, 2003; Nittrouer, 2002; Nittrouer & Miller, 1997; Tyler & Saxman, 1991). Children with SSD and high speech error variability may have delayed development of these representations.

Speech error variability was also negatively correlated with $MLU_m$ and the Core Language score from the CELF–P2. That is, children with the most variable speech errors showed the poorest language abilities. Broomfield and Dodd (2004) also found the poorest language skills in children with the most variable speech output, although their study focused on word variability and not speech error variability. As alluded to by Preston and Koenig (2011), the relationship between language skills and variable output may be bidirectional: Variable output may hinder a child's ability to map linguistic information to a spoken production. The converse may also be true: Unstable linguistic representations may lead to variable output.

## Limitations

There are several limitations associated with the present study. First, the two measures of speech variability showed high variability among the participants. For example, for word variability, the standard deviation was approximately half of the mean value. This high variability is likely due to the small sample size and may be at least partially responsible for the nonsignificant correlations. Given the

small sample size and the nonsignificant correlations, the findings from the present study should be considered preliminary. Second, the participants presented with a wide range of ages (3;6–5;5). It is not unreasonable to assume that any relationship or lack of relationship seen between different measures of speech variability and between these measures and other measures of speech and language may be dependent on the age of the speaker. Recall that Preston and Koenig (2011) found high correlations between measures of word and speech error variability in their participants. There was no such relationship in the present study's participants, who were younger than those in Preston and Koenig. In order to obtain the clearest picture of lexical- and phonological-level processing at any point in development, a sufficiently narrow view of development is necessary. The wide range of ages of the participants also may have been partially responsible for the nonsignificant correlations. Third, although the omission of vowels from calculations of word variability is not without precedent (Ferguson & Farwell, 1975; Ingram, 2002; Leonard, Rowan, Morris, & Fey, 1982; Sosa & Stoel-Gammon, 2006, 2012), it is considered a limitation in the present study. Including vowels in the calculations would lead to more sensitive measures of word variability, particularly for children with vowel difficulties. This is especially important when studying children with suspected or diagnosed CAS given that such a diagnosis is based, in part, on "inconsistent errors on consonants and vowels" (ASHA, 2007, p. 2). Future research in word variability in children should include vowels in variability calculations.

## Clinical Implications

In preschool children with SSD, it would appear that high levels of speech variability reflect underlying linguistic deficits. These findings have important clinical implications. First, preschool children with SSD and high word variability should be monitored closely to ensure that they are acquiring vocabulary knowledge at a rate that is typical for their age. Clinicians should aggressively target vocabulary expansion in children with SSD and vocabulary deficits. Second, according to Metsala and Walley's (1998) LRM, a delay in the segmental restructuring of lexical representations can lead to reading disabilities in some children. Given the present study's finding of an inverse relationship between speech error variability and SRT performance, it would seem that children with high speech error variability are at risk for reading disabilities because of indistinct phonological representations. Clinicians should monitor these children for potential phonemic awareness and reading deficits and remediate these when they are present. Identifying underlying deficits in children with SSD will allow us to design more focused interventions that target these deficits. In children with SSD and high speech variability, interventions that address not just the accuracy of production but also its consistency may increase the stability of their underlying representations and lead to the greatest treatment gains.

## References

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Position statement]. Available from www.asha.org/policy

Bankson, N., & Bernthal, J. E. (1990). *Bankson–Bernthal Test of Phonology*. Austin, TX: Pro-Ed.

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37,* 391–403.

Bradford, A., & Dodd, B. (1994). The motor planning abilities of phonologically disordered children. *European Journal of Disorders of Communication, 29,* 349–369.

Broomfield, J., & Dodd, B. (2004). The nature of referred subtypes of primary speech disability. *Child Language Teaching and Therapy, 20,* 135–151.

Burt, L., Holm, A., & Dodd, B. (1999). Phonological awareness skills of 4-year-old British children: An assessment and developmental data. *International Journal of Language and Communication Disorders, 34,* 311–335.

Coady, J. A., Evans, J. L., Mainela-Arnold, E., & Kluender, K. R. (2007). Children with specific language impairment perceive speech most categorically when tokens are natural and meaningful. *Journal of Speech, Language, and Hearing Research, 50,* 41–57.

dePaola, T. (1978). *Pancakes for breakfast*. Orlando, FL: Harcourt.

Dodd, B. (1995). *Differential diagnosis and treatment of children with speech disorder*. London, United Kingdom: Whurr.

Dodd, B. (2005). *Differential diagnosis and treatment of children with speech disorder* (2nd ed.). London, United Kingdom: Whurr.

Dodd, B., Holm, A., Crosbie, S., & McIntosh, B. (2010). Core vocabulary intervention. In A. Williams, S. McLeod, & R. J. McCauley (Eds.), *Interventions for speech sound disorders in children* (pp. 117–136). Baltimore, MD: Brookes.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Dunn, L., Dunn, L., & Williams, K. T. (1997). *Peabody Picture Vocabulary Test—III*. Circle Pines, MN: AGS.

Elbro, C., Borstrom, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33,* 36–60.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51,* 419–439.

Forrest, K., Dinnsen, D. A., & Elbert, M. (1997). Impact of substitution patterns on phonological learning by misarticulating children. *Clinical Linguistics & Phonetics, 11,* 63–76.

Forrest, K., Elbert, M., & Dinnsen, D. A. (2000). The effect of substitution patterns on phonological treatment outcomes. *Clinical Linguistics & Phonetics, 14,* 519–531.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27,* 513–543.

Guilford, J. P. (1956). *Fundamental statistics in psychology and education*. New York, NY: McGraw–Hill.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics, 28,* 377–396.

Holm, A., Crosbie, S., & Dodd, B. (2005). Treating inconsistent speech disorders. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorder* (pp. 182–201). London, United Kingdom: Whurr.

Holm, A., Crosbie, S., & Dodd, B. (2007). Differentiating normal variability from inconsistency in children's speech: Normative data. *International Journal of Language and Communication Disorders, 42,* 467–486.

Holm, A., Farrier, F., & Dodd, B. (2008). Phonological awareness, reading accuracy and spelling ability of children with inconsistent phonological disorder. *International Journal of Language and Communication Disorders, 43,* 300–322.

Ingram, D. (2002). The measurement of whole-word productions. *Journal of Child Language, 29,* 713–733.

Iuzzini, J., & Forrest, K. (2011, November). *Lexical and segmental inconsistency in disordered and typical developing speech.* Poster session presented at the annual convention of the American Speech-Language-Hearing Association, San Diego, CA.

Kim, H., Hasegawa-Johnson, M., & Perlman, A. (2010). Vowel contrast and speech intelligibility in dysarthria. *Folia Phoniatrica et Logopaedica, 63,* 187–194.

Leonard, L. B., Rowan, L. E., Morris, B., & Fey, M. E. (1982). Intra-word phonological variability in young children. *Journal of Child Language, 9,* 55–69.

Liker, M., & Gibbon, F. E. (2008). Tongue palate contact patterns of velar stops in normal adult English speakers. *Clinical Linguistics & Phonetics, 22,* 137–148.

Marquardt, T. P., Jacks, A., & Davis, B. L. (2004). Token-to-token variability in developmental apraxia of speech: Three longitudinal case studies. *Clinical Linguistics & Phonetics, 18,* 127–144.

Mayo, C., Scobbie, J. M., Hewlett, N., & Waters, D. (2003). The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. *Journal of Speech, Language, and Hearing Research, 46,* 1184–1196.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.

Miller, J. F., & Chapman, R. S. (2000). Systematic Analysis of Language Transcripts (Version 6.1) [Computer software]. Madison, WI: Waisman Center.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48,* 1033–1047.

Nittrouer, S. (2002). From ear to cortex: A perspective on what clinicians need to understand about speech perception and language processing. *Language, Speech, and Hearing Services in Schools, 33,* 237–252.

Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *The Journal of the Acoustical Society of America, 101,* 2253–2266.

Preston, J. L., & Koenig, L. L. (2011). Phonetic variability in residual speech sound disorders. *Topics in Language Disorders, 31,* 168–184.

Rispens, J., & Baker, A. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research, 55,* 683–694.

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

Semel, E., Wiig, E., & Secord, W. (2004). *Clinical Evaluation of Language Fundamentals—Preschool, Second Edition*. San Antonio, TX: The Psychological Corporation.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders: III. A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47,* 256–270.

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009). A nonword repetition task for speakers with misarticulations: The Syllable Repetition Task (SRT). *Journal of Speech, Language, and Hearing Research, 52,* 1189–1212.

Sosa, A. V., & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language, 33,* 31–50.

Sosa, A. V., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language, and Hearing Research, 55,* 596–608.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language, 38,* 1–34.

Thelen, E., & Bates, E. (2003). Connectionism and dynamic systems: Are they really different? *Developmental Science, 6,* 378–391.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Tyler, A. A., & Lewis, K. E. (2005). Relationships among consistency/variability and other phonological measures over time. *Topics in Language Disorders, 25,* 243–253.

Tyler, A. A., Lewis, K. E., & Welch, C. M. (2003). Predictors of phonological change following intervention. *American Journal of Speech-Language Pathology, 12,* 289–298.

Tyler, A. A., & Saxman, J. H. (1991). Initial voicing contrast acquisition in normal and phonologically disordered children. *Applied Psycholinguistics, 12,* 453–479.

van Geert, P., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior & Development, 25,* 340–374.

Williams, P., & Stephens, H. (2010). The Nuffield Centre Dyspraxia Programme. In A. Williams, S. McLeod, & R. J. McCauley (Eds.), *Interventions for speech sound disorders in children* (pp. 159–177). Baltimore, MD: Brookes.

Ziegler, W., Hartmann, E., & Hoole, P. (1993). Syllabic timing in dysarthria. *Journal of Speech and Hearing Research, 36,* 683–693.