

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Special Education and Communication  
Disorders Faculty Publications

Department of Special Education and  
Communication Disorders

---

2009

## A Nonword Repetition Task for Speakers with Misarticulations: The Syllable Repetition Task (SRT)

Lawrence D. Shriberg  
*University of Wisconsin-Madison*

Heather L. Lohmeier  
*University of Wisconsin-Madison*

Thomas F. Campbell  
*University of Texas-Dallas, tfc061000@utdallas.edu*

Christine A. Dollaghan  
*University of Texas-Dallas*

Jordan R. Green  
*University of Nebraska-Lincoln, jgreen4@unl.edu*

*See next page for additional authors*

Follow this and additional works at: <https://digitalcommons.unl.edu/specedfacpub>



Part of the [Special Education and Teaching Commons](#)

---

Shriberg, Lawrence D.; Lohmeier, Heather L.; Campbell, Thomas F.; Dollaghan, Christine A.; Green, Jordan R.; and Moore, Christopher A., "A Nonword Repetition Task for Speakers with Misarticulations: The Syllable Repetition Task (SRT)" (2009). *Special Education and Communication Disorders Faculty Publications*. 10.

<https://digitalcommons.unl.edu/specedfacpub/10>

This Article is brought to you for free and open access by the Department of Special Education and Communication Disorders at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Special Education and Communication Disorders Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

---

**Authors**

Lawrence D. Shriberg, Heather L. Lohmeier, Thomas F. Campbell, Christine A. Dollaghan, Jordan R. Green, and Christopher A. Moore

A Nonword Repetition Task  
for Speakers with Misarticulations: The Syllable Repetition Task (SRT)

Lawrence D. Shriberg<sup>a</sup>

Heather L. Lohmeier<sup>a</sup>

Thomas F. Campbell<sup>b</sup>

Christine A. Dollaghan<sup>b</sup>

Jordan R. Green<sup>c</sup>

Christopher A. Moore<sup>d</sup>

<sup>a</sup>Waisman Center, University of Wisconsin-Madison

<sup>b</sup>The University of Texas-Dallas

<sup>c</sup>University of Nebraska-Lincoln

<sup>d</sup>National Institute on Deafness and Other Communication Disorders

Address correspondence to:

Lawrence D. Shriberg  
Room 439, Waisman Center  
University of Wisconsin-Madison  
1500 Highland Avenue  
Madison, WI 53705

This is an author-produced manuscript that has been peer reviewed and accepted for publication in the *Journal of Speech, Language, and Hearing Research (JSLHR)*. As the "Papers in Press" version of the manuscript, it has not yet undergone copyediting, proofreading, or other quality controls associated with final published articles. As the publisher and copyright holder, the American Speech-Language-Hearing Association (ASHA) disclaims any liability resulting from use of inaccurate or misleading data or information contained herein. Further, the authors have disclosed that permission has been obtained for use of any copyrighted material and that, if applicable, conflicts of interest have been noted in the manuscript.

### Key Words

assessment, genetics, memory, speech disorders, endophenotype

### Running Head

#### Syllable Repetition Task (SRT)

**Purpose.** Conceptual and methodological confounds occur when non(sense) repetition tasks are administered to speakers who do not have the target speech sounds in their phonetic inventories or who habitually misarticulate targeted speech sounds. We describe a nonword repetition task, the *Syllable Repetition Task (SRT)* that eliminates this confound and report findings from three validity studies.

**Method.** Ninety-five preschool children with Speech Delay and 63 with Typical Speech, completed an assessment battery that included the *Nonword Repetition Task (NRT)* (Dollaghan & Campbell, 1998) and the SRT. SRT stimuli include only four of the earliest occurring consonants and one early occurring vowel.

**Results.** Study 1 findings indicated that the SRT eliminated the speech confound in nonword testing with speakers who misarticulate. Study 2 findings indicated that the accuracy of the SRT to identify expressive language impairment was comparable to findings for the NRT. Study 3 findings illustrated the SRT's potential to interrogate speech processing constraints underlying poor nonword repetition accuracy. Results supported both memorial and auditory-perceptual encoding constraints underlying nonword repetition errors in children with speech-language impairment.

Conclusion. The SRT appears to be a psychometrically stable and substantively informative nonword repetition task for emerging genetic and other research with speakers who misarticulate.

### *Nonword Repetition Tasks in Genetic Studies of Verbal Trait Disorders*

Findings from the genetics literature continue to support Adams and Gathercole's (2000) conclusion that poor non(sense)word repetition is a key feature of heritable specific language impairment (SLI). Bishop and colleagues (Bishop, 2002a, 2002b; Bishop, Adams, & Norbury, 2004; Bishop, Bishop, Bright, Delaney, & Tallal, 1999; Bishop, North, & Donlan, 1996), and Kovas et al. (2005) have reported high heritability for nonword repetition in twin samples and cohorts of children with concurrent or histories of speech-language disorders. Shriberg et al. (2005) reported that in comparison to control participants with speech sound disorders, participants with speech sound disorders who were at familial risk for a genetically transmitted subtype of speech disorder had significantly lower nonword repetition task performance. Molecular genetic studies using nonword repetition tasks have reported linkage to regions of interest on chromosomes 16 and 19 for children with language impairment (Monaco and the SLI Consortium [SLIC], 2007; SLI Consortium, 2002, 2004). The first genetic entries for speech sound disorder in the Online Mendelian Inheritance in Man database were based, in part, on linkage of nonword repetition task performance to regions of interest on chromosome 3 (Stein et al., 2004) and chromosome 6 (Smith, Pennington, Boada, & Shriberg, 2005; see Caylak, 2007, and Lewis et al., 2006 for literature reviews).

One property that appears to underlie the productivity of nonword repetition tasks in genetic and other research is their sensitivity to a range of behavioral *phenotypes* (i.e., characteristics used to classify and quantify traits and disorders). Measures that are sensitive to but not specific for a target phenotype, termed *endophenotypes* (Gottesman & Gould, 2003), may be so because they are more closely influenced by gene products than are phenotypes. Accordingly, the *heritability* (variance in a disorder associated with genetic rather than

environmental sources; but see Butcher & Plomin, 2008; Plomin & Davis, 2006) of endophenotypes may be higher than the heritability of a disorder studied using a phenotype specific for the disorder, particularly when the disorder may be genetically heterogeneous.

Another likely reason for the robust findings for nonword repetition tasks as endophenotypes in genetic and other studies of verbal traits and disorders is that success on such tasks is dependent on the integrity of speech processing at several pre-execution stages (see Ellis Weismer & Edwards, 2006 for a discussion of interactivity among such putatively serial constructs). As conceptualized in diverse neurocognitive, psycholinguistic, and speech motor control frameworks (e.g., Bock, 1982; Dell, 1986; Guenther, 1995; Levelt, 1989; Stackhouse & Wells, 1997; van der Merwe, in press; Ziegler, 2006) these perspectives generally include: (a) auditory-perceptual processes that allow veridical encoding of phonemic, sublexical, and lexical representations of the stimulus to be repeated; (b) memorial processes that store and retrieve these representations; and (c) planning/programming processes that configure the movement gestures preceding manifest speech. Specific mechanisms imputed for each of these three components of speech processing, as well as the type and relative influence of mediating and moderating variables (e.g., age, gender, phonological awareness, processing speed, articulation rate), differ considerably within and among the many disciplines that use nonword repetition tasks to study verbal behavior and its disorders.

In addition to their use as endophenotypes to identify persons at risk for, with concurrent, or with resolved verbal trait disorders in genetic studies, the more widely reported uses of nonword repetition tasks are to inform theories of lexical acquisition and to identify language disorders in speakers from diverse cultural communities without the confounds in language assessment associated with such variables as vocabulary, world knowledge, and dialect (e.g.,

Campbell, Dollaghan, Needleman, & Janosky, 1997; Ellis Weismer et al., 2000; Oetting & Cleveland, 2006; Rodekohr & Haynes, 2001; Washington & Craig, 2004). Two recent comprehensive reviews of the extensive nonword repetition task literature summarize effect sizes and likelihood ratios quantifying the diagnostic accuracy of nonword repetition tasks to identify language impairment (Coady & Evans, 2008; Graf Estes, Evans, & Else-Quest, 2007) and many papers have reviewed associated methodological and substantive issues (e.g., Archibald & Gathercole, 2006; Sahlen, Reuterskiöld-Wagner, Nettelbladt, & Radeborg, 1999). The following discussion focuses on issues motivating development of the Syllable Repetition Task and the questions posed in three studies.

#### *Nonword Repetition Tasks in Speech-Language Research*

A continuing question in nonword repetition task performance is the proportion of variance in speech-language impairment accounted for by memorial processes, compared to contributions from other sources of variance in nonword repetition scores as proxies for processing constraints in lexical acquisition. Although the extensive theoretical and clinical contributions of this conception continue to support the original and updated proposal that poor nonword repetition task accuracy indicates a deficit in a specific aggregate of memorial processes (Baddeley, 2000, 2003; Baddeley & Hitch, 1974; Baddeley & Logie, 1999), and that such deficits underlie language and other verbal trait impairment, there is extensive support for the perspective that real and nonword learning is multiply determined. Archibald and Gathercole (2006) conclude that findings from an increasing number of nonword repetition studies “. . . cannot be readily accommodated by a verbal short-term memory deficit account of specific language impairment” (p. 979). Findings from a series of well-controlled word learning studies assessing auditory-perceptual processes by Edwards and Munson and colleagues (e.g., Edwards,

Fox, & Rogers, 2002; Munson, 2006; Munson, Edwards, & Beckman, 2005; Munson, Kurtz, & Windsor, 2005) have led these investigators to conclude that children with speech sound disorder have “poorly specified primary representations” due to “difficulties forming robust representations of the acoustic-auditory and articulatory characteristics of speech” (Munson, Baylis, Krause, & Kim, in press). Munson and colleagues (in press) further posit that auditory-perceptual encoding deficits may be the source of speech production errors in speech sound disorders, noting that many speech production models emphasize that “. . . acoustic/perceptual representations for words serve as targets during speech production (e.g., Guenther, 1995; Pierrehumbert, 2002).” Relevant findings and discussions on the contribution of auditory-perceptual encoding processes to word learning and to the accuracy of nonword repetition include Alphen et al., 2004; Bishop et al., 1999; Coady, Evans, Mainela-Arnold, and Kluender, 2006; Coady, Kluender, and Evans, 2005; Evans, Viele, Kass, and Tang, 2002; Reuterskiöld-Wagner, Sahlén, and Nyman, 2005; Rvachew, Ohberg, Grawburg, and Heyding, 2003; and Storkel, 2004.

Support for the contribution of articulatory planning/programming as a third source of inaccurate nonsense word repetition is best characterized as equivocal. A primary interpretive constraint is the wide range of constructs, pathway models, and assessment approaches that may be invoked to assess speech motor competence (for a complex performance example, see Howard, Binks, Moore, & Playfer, 2000). From the speech acquisition literature, Sahlén et al. (1999) reported that children with speech disorder had poor performance on a nonword repetition task, but participants’ accuracy scores were not significantly associated with their scores on a brief task assessing oral motor skills. Stark and Blackwell (1997) reported significant associations between nonword repetition scores and accuracy/coordination scores on a series of

oral movement tasks for children with language and speech impairment. Edwards and Lahey (1998) reported equivocal findings from an error analysis of nonword repetition responses comparing accuracy percentages on relatively more difficult articulatory motor targets (fricatives, liquids, clusters) to scores on less challenging targets for children with SLI and controls. However, both Bishop et al. (1996) and Briscoe, Bishop, and Norbury (2001) have reported significant influences of articulatory complexity on the nonword repetition scores of children with SLI. Adams and Gathercole (2000) found that for groups of typically-speaking children with good and poor nonword repetition performance, the language/memory relationship was not eliminated when the output requirements of the nonword repetition task were minimized by requiring pointing rather than speech responses. Analyses and syntheses of these and other findings are available in Coady and Evans' (2008) comprehensive review of speech processing constraints on nonword repetition task accuracy and in Gathercole's (2006) keynote paper and follow-up comments from 14 investigator groups.

#### *Nonword Repetition Tasks to Identify Language Impairment*

Graf Estes et al. (2007) provide a comprehensive history of nonword repetition task research in studies of children and adults with typical and atypical language acquisition. As suggested by the increasing number of reports of new nonword repetition tasks and language adaptations internationally of extant measures, they are becoming standard components in assessment protocols to identify language disorder. Extensive theoretical and methodological discussion in Graf Estes et al. (2007) concerns the variance in nonword repetition accuracy associated with the most commonly used nonword tasks used in research and with characteristics of the stimulus in these tasks such as their wordlikeness, perceptual salience, predictability, frequency of occurrence, phonotactic frequencies, stress assignment, density of phonological

neighborhood, and articulatory complexity (the latter variable is discussed in detail in the following section).

A major evidenced-based need in the use of nonword repetition tasks to identify language impairment are likelihood ratios estimating a nonword repetition task's accuracy in identifying concurrent or risk for future verbal impairment. The meta-analytic and moderator analyses of 23 studies reported in Graf Estes et al. (2007) indicated that although nonword repetition accuracy for children with SLI averaged 1.3 standard deviations below the performance of children with typical language across all measures, nonword repetition tasks cannot presently be considered interchangeable for research and clinical use.

Ellis Weismer et al. (2000) reported one of the largest set of findings on this question based on responses to the Nonword Repetition Task (NRT: Dollaghan & Campbell, 1998). These investigators provided likelihood ratios estimating the diagnostic accuracy of the NRT to identify both specific and nonspecific language impairment in a database of 581 second grade children. Participants had been identified using population-based procedures (Tomblin et al., 1997) and their language status in a longitudinal study was classified each year as specific or nonspecific language impairment using a diagnostic algorithm that has since been widely applied (Tomblin, Records, & Zhang, 1996). Ellis Weismer and colleagues reported positive likelihood ratios of nearly 3.0 indicating that children with scores below an optimized cut point on the NRT were at nearly three times the risk for specific and nonspecific language disorder (see Ellis Weismer et al., 2000, Table 3 and accompanying text for relevant detail). Citing Sackett, Haynes, Guyatt, and Tugwell's (1991) quantitative criteria for the clinical use of likelihood ratios, the obtained diagnostic accuracy values were termed 'intermediate.' Ellis Weismer and colleagues cautioned that although nonword repetition tasks can aid in identifying children who perform poorly on

standard language measures, scores on the nonword task used in their study were not sufficient in themselves to classify language disorder for children in this age range, although they accurately identified children who were currently enrolled in language treatment.

#### *A Measurement Confound in Nonword Repetition Tasks*

Whether used as an endophenotype in genetic studies, to inform theories of language development and disorder, or to identify speakers with language disorder, a measurement confound occurs when nonword repetition tasks are administered to speakers who misarticulate speech sounds. Such speakers include young, typically developing children with incomplete phonetic inventories, preschool and older children with speech delays of unknown origin, and children and adults of any age with a speech sound disorder (SSD) associated with impairments in cognitive, sensory, motor, structural, or affective processes of known origin. The speech confound may occur for the speaker, the examiner, or both. Speakers may have difficulty discriminating and/or encoding phonemes, storing and/or retrieving representations, or planning gestures and/or programming movements for speech sounds in nonwords that are not in their phonetic inventory — or that are in their phonetic inventory, but have not been fully mastered (Bishop & Edmundson, 1987; Briscoe et al., 2001; Hill, 2001; Sahlen et al., 1999). The importance of this theoretical and methodological association was noted over a decade and a half ago by pioneers in this area. In a study in which phonological working memory findings were interpreted as reflecting constraints in speech rate, Gathercole and Adams (1993) concluded that “. . . phonological memory skills are intimately linked with many aspects of speech production in early childhood” (p. 777).

For the examiner, speech production errors affecting the precision or intelligibility of responses pose psychometric challenges to the reliability and validity of nonword repetition task

scores. First, examiners must have reliable transcription skills and a set of well-developed conventions to score repetition errors. For example, most nonword tasks instruct the examiner to score distortions (e.g., a dentalized /s/ or a derhotacized /r/) as a correct repetition of a speech sound, but score substitutions (e.g., /ʃ/ for /s/ or /w/ for /r/) as repetition errors. The ability to distinguish distortions from substitutions requires training and typically has a large standard error of measurement (Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997a). Transcription instructions seldom address other technical considerations, such as whether devoiced voiced obstruents are to be scored as correct or incorrect. Especially when assessing speakers with significant speech impairment, including those in emerging genetic studies of speakers with motor speech disorders (i.e., apraxia of speech, dysarthria), the standard error of measurement on nonword repetition tasks is significantly affected by the extensiveness of the transcription rules and the reliability of one or more examiner's transcription skills.

A second problem for examiners is the classification of nonword repetition errors: are all errors to be scored as repetition errors or could some possibly be “explained” by a speaker's habitual misarticulation of the target phoneme. Thal, Miller, Carlson, and Moreno Vega (2005) concluded that misarticulations were not a confound in nonword repetition task scores obtained from 4-year-old children. This appropriate conclusion was based on findings from a study of children who were in the normal range on a number of receptive and expressive tests of language because such children were less likely to have SSD. However, review of widely cited studies using nonword repetition tasks, both within and beyond the linguistic-genetics literature, indicates that this second type of potential speech confound is a scoring problem that has been addressed in four ways.

As most nonword repetition studies are not focused on children or adults with significant misarticulations, the most frequent scoring approach has been to ignore the possibility that errors in nonword repetition could reflect habitual speech production errors. Procedures to transcribe and score misarticulations may not be deemed necessary because, as noted in Thal et al. (2005), in some studies they are infrequent. An alternative rationale, as proposed in one speech-genetics study, is to score misarticulations in nonword tasks as incorrect because all errors in speech sound production, whatever their source, are included in the broad phenotype under study (Kovas et al., 2005; Stein et al., 2004).

A second approach to address the speech confound in nonword repetition task responses is to eliminate from analyses participants with SSD (e.g., Archibald & Gathercole, 2006). Bishop et al. (1996) and Ellis Weismer et al. (2000) excluded from their analyses of nonword repetition task data, all participants who scored below 85% accuracy, on a citation-form articulation test or a conversational sample, respectively. This approach is not feasible for study samples predicted to have a high prevalence of SSD, or more pointedly, samples in which active SSD is the primary phenotype under study.

A third and frequently employed approach in scoring nonword repetition tasks is to give speakers some type of scoring credit for sounds that meet investigator-determined criteria for being misarticulated elsewhere in the available speech samples (e.g., from responses to one or more articulation tests or as produced in a conversational speech sample). For example, in their study of children with language impairment, Edwards and Lahey (1998) noted “. . . if a child evidenced a consistent substitution or distortion in spontaneous speech (e.g., a /ʃ/ for /s/ substitution or a frontal lisp), then this substitution was not scored as incorrect” (p. 288). As noted previously, such procedures require that additional speech samples are available to identify

the misarticulation, that the investigators have the requisite perceptual skills to identify and transcribe speech errors reliably, and that standardized procedures are in place to determine the error types, positions in words, and numbers of tokens needed to attest each consonant in singleton and cluster contexts and each vowel/diphthong as ‘in’ versus ‘not in’ a child’s phonetic inventory. All such technical decisions, for which there are no well-established conventions, may contribute to variance in the standard error of measurement of nonword repetition task scores.

A fourth approach to the speech production confound in nonword repetition tasks has been to restrict the speech sounds in the task stimuli to those that most children do not misarticulate (e.g., Montgomery, 1995). Dollaghan and Campbell (1998) used this and associated rationale in developing the *Nonword Repetition Task* (NRT). They excluded from the 16 NRT nonword stimuli, the 8 English consonants (termed the *Late-8*; Shriberg, 1993) that are mastered latest in children with both typical and atypical speech acquisition. The appendix provides a list of the NRT stimuli, which include 9 different vowels and diphthongs and 11 different consonants for a total of 20 different speech sounds. We note that younger children with typical speech and children with more significantly involved speech may misarticulate one or more of the 5 vowels and especially the 4 diphthongs (Pollack & Berni, 2003; Pollack & Keiser, 1990; Stoel-Gammon & Herrington, 1990) and that children with SSD will likely misarticulate one or more of the 11 consonant sounds in singleton environments (Shriberg & Kwiatkowski, 1994).

### *Research Questions*

The Syllable Repetition Task (SRT) to be described is a nonword repetition task developed for use in genetic and other studies that include young speakers with limited phonetic inventories or speakers of any age with speech sound disorders of known or unknown etiology. The primary goal of this nonword repetition task is to provide a means to examine speech

processing constraints, while minimizing or eliminating speaker, scoring, and interpretive confounds associated with misarticulations. In the following section, we discuss the development of the SRT, describe the methods used to collect and analyze SRT and NRT data from a sample of 158 participants, and summarize the primary psychometric findings. We then report results from a proof of concept study, a study of the accuracy of the SRT to identify children with expressive language impairment, and a study of potential speech processing constraints underlying poor nonword repetition accuracy in young children with speech delay and expressive language impairment. The research questions posed in each of the three studies, respectively, are:

- 1. Does the SRT eliminate the speech confound when assessing speakers with misarticulations?*
- 2. Does the SRT identify children with expressive language impairment?*
- 3. Does the SRT provide information on speech processing constraints that may underlie poor nonword repetition task performance?*

## Method

### *Construction of the SRT*

#### *Preliminary Studies*

Preliminary studies were completed with several dozen children attending a university speech clinic to explore the potential of a phoneme-limited approach to nonword repetition task testing for young children and speakers of any age with significant speech impairment. These studies suggested that a small set of acoustically salient (i.e., voiced) anterior consonants might be both available in the inventories of young speakers and those with a speech impairment and be intelligible to an examiner, regardless of the precision with which they were repeated. These

pilot studies suggested that four consonants within two manner features (stops, nasals) best met these criteria: /b/, /d/, /m/, and /n/.

A computer-based analysis was then completed on audio-recorded conversational speech samples from 268 children (184 males and 84 females; mean age = 4;5 years; range = 3;0-5;11 years) who had participated in research studies with the Phonology Project & Clinic, Waisman Center, University of Wisconsin-Madison. The samples included children whose speech impairments reflected the full range of speech delay of unknown origin. The inclusionary restrictions placed on this search were that the samples had adequate recording quality, were sufficient in length, and had been transcribed by research assistants with documented transcription reliability. This search indicated that 100% of the children, including those with *moderate-severe* and *severe* speech sound disorder (i.e., Percentage of Consonants Correct scores lower than 65%), had at least one instance of /b/, /m/, /n/ and /e/ in their phonetic inventories, and all but two children (99.3%) had at least one correct production of /d/.

Following trials of several pilot versions, the content and administration procedures for an 18-item nonword task termed the *Syllable Repetition Task (SRT)* were finalized. The appendix lists the stimuli for the SRT, which include eight 2-syllable (CVCV) items (e.g., *dama*), six 3-syllable (CVCVCV) items (e.g., *nabada*), and four 4-syllable (CVCVCVCV) items (e.g., *manadaba*). The number of items at each level of difficulty was motivated by the target groups for which the SRT was developed, with proportionally more ‘easy’ items positioned first for children with cognitive, linguistic, and/or motivational challenges. The stressed vowel /e/ occurs in each syllable. Additional considerations were that the two consonants in each manner class were balanced as best as possible in their distribution within words and across syllable levels.

An adult female was trained to clearly and naturally speak each nonword with equal stress on each of the constituent syllables, thus optimizing for both the speaker and the examiner/transcriber, the acoustic cues in the /e/ vowel for each initial and intervocalic consonant. After a series of pilot studies in which the stimuli were presented on analog tapes, a digital version was created for laptop presentation.

The appendix includes the website address from which the SRT Technical Report and all task materials can be downloaded without cost (Shriberg & Lohmeier, 2008). The SRT Technical Report includes (a) tabular data and text discussion of psychometric findings, (b) tabular data and discussion of statistical findings from three procedural analyses, (c) tabular comparison data obtained from 70 children ages 4-to-16 years with typical speech, (d) administration instructions, (e) scoring instructions, and (f) a form for manual scoring of the SRT

### *Rationale*

The SRT stimuli listed in the appendix are claimed to have the following seven properties: (a) the four consonants and one vowel will be in the phonetic inventories of young children and all but the most severe speakers with speech sound disorders; (b) the three simple syllabic structures (CVCV, CVCVCV, CVCVCVCV) eliminate opportunities for respondents to delete final consonants or to reduce clusters; (c) respondents will perceive these syllable trains as potential words, even without the presence of stress cues (Bonnot, 1999; Gallon, Harris, & van der Lely, 2007; Marton, 2006; Roy & Chiat, 2004); (d) the task demands of repeating only four phonemes correctly, but in ‘words’ as long as 4-syllables will provide a sufficient challenge to the speech processing mechanisms proposed to be invoked when responding to nonword repetition tasks stimuli; (e) the relatively short time needed to complete the task will minimize the risk of fatigue in nonword repetition tasks that have included as many as 96 items (Gallon et

al., 2007); (f) distorted, but phonemically correct productions of the four consonants (e.g., as may occur in responses of speakers with subtypes of dysarthria) will readily be discriminable for examiners with minimal skills in phonetic transcription; and (g) the inclusion of two consonants within two manner classes (i.e., two nasals, two voiced stops) provides a means for error analyses of within- versus between-class consonant substitutions (to be described). On the latter claim, the structural and phonemic features of the SRT stimuli described in (a) through (g) are proposed as conceptual and methodologic requirements for valid and reliable quantitative error analyses of responses to the SRT to be described. One claim that cannot be made about the stimuli is that none of the constituent syllables or items in the SRT are real words (consider ‘ma’ in English and ‘nada’ in Spanish). However, as claimed in (c) above, the equal stress on each syllable (e.g., ‘nede’), which was designed to maximize auditory information on the formant loci for the consonants and to remove the contribution of prosodic variables to repetition scores, attenuates the wordlikeness of the stimuli.

### *Scoring*

Procedures used to score SRT responses are similar to those used to score the NRT. The SRT Technical Report described in the appendix provides complete information on manual scoring of the SRT. Briefly here, deletions and substitutions of sounds are scored as incorrect, including both consonants and vowels/diphthongs in the NRT but only the four consonants in the SRT. Speech sound distortions and possible cognate substitutions on the voiced stops (i.e., p/b, t/d) are scored as correct and addition errors are ignored, provided the correct phoneme occurs in the correct position. Rather than whole item scoring, which does not allow for partial credit, correct repetition of each of the 50 consonant targets in the SRT contributes two points to the total score.

Preliminary statistical analyses were completed to determine rationale and procedures to score non-responses. Details of the analyses are included in the Technical Report. Essentially, based on the 158 participants' scores computed using three scoring systems, findings from this study motivated the decision to score all non-responses to SRT items as errors, provided the examiner viewed the administration as valid. As in administration guidelines for all formal testing, the recommendation is to retest children whose responses to the SRT are suspected to be invalid for any reason (e.g., participant was tired, not attending, etc.).

*Database for Psychometric Studies and Three Construct Validity Studies of the SRT*

*Participants*

Data from 179 participants in a physiology study of typical and atypical speech development (Moore et al., 2006) were screened for potential use in the present study series. Participants in this database had completed an assessment protocol that included the NRT and the SRT. Children with speech delay (SD) of unknown origin (Shriberg et al., 2005) had been recruited from referrals to the Department of Audiology and Communication Disorders at the Children's Hospital of Pittsburgh. Participants with typical speech (TS) and no history of speech, language, or other developmental delays had been recruited through childcare and preschool centers in the Pittsburgh area. Children in the four typical speech groups were selected from the available pool to be as similar as possible to the participants with SD in age, and secondarily on sex, maternal education, and maternal ethnicity. All children came from homes in which English was the only spoken language.

As described below, conversational speech samples available in the assessment protocol for the physiology study had been obtained by examiners using sampling conventions developed for research in SSD (Shriberg & Kwiatkowski, 1980, 1983, 1985). The samples were analyzed

and classified using the Speech Disorders Classification System (SDCS: Shriberg, 1993, 2008; Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997b), a program in the PEPPER software suite (Shriberg, Allen, McSweeney, & Wilson, 2001). A total of 21 of the 179 eligible children were excluded from the present study series because they did not clearly meet the SDCS classification criteria for TS (16 children) or SD (5 children). The remaining group of 158 participants included 63 children with TS and 95 with SD. Complete SRT responses were available for all participants with TS and 93 of the 95 participants with SD.

To provide developmental information on the SRT, participants in the TS and SD groups were subdivided into the four age groups shown in Table 1. As indicated in Table 1, comparable values for sex, maternal education, and maternal ethnicity across the eight subgroups were not obtained for the TS and SD participants whose data were eligible for the present study. TS and SD participants differed in sex (SD participants were proportionally more male;  $\chi^2 = 15.20$ ;  $p = .001$ ), and SD participants' mean maternal education was lower ( $t = 2.66$ ;  $p = .009$ ). At the subgroup level, the TS and SD participants differed in age (SD participants were older in Group 3 [ $t = 2.76$ ;  $p = .009$ ] and Group 4 [ $t = -9.06$ ;  $p = .001$ ]), sex (Group 1 had a higher percentage of boys [ $\chi^2 = 10.45$ ;  $p = .001$ ]) and maternal education (Group 1 had higher levels of maternal education [ $t = 2.77$ ;  $p = .008$ ]). Finally, as shown in Table 1, participants came from ethnically diverse backgrounds. A test of proportions indicated that the aggregated percentage of non-European Americans in the TS and SD groups did not differ at the total group level ( $Z = 1.58$ ;  $p = .113$ ), but there were substantial differences in the percentages of participants from each background across subgroups. These group and subgroup level differences across several sociodemographic variables should be viewed as possible constraints on generalizations from findings.

### *Assessment*

Speech, language, and hearing assessments in the original study were completed by the examiners over a two-day period. None of the participants had known structural, neurological, or affective deficits, as documented by caregivers' responses on a case history form and reports from previous speech-language services. Hearing was screened bilaterally at 1, 2, and 4 kHz using earphones. Participants with thresholds above 25 dB HL for more than one frequency in the same ear were brought back at another date to complete the assessment protocol. One child failed follow-up audiological screening and was excluded from the study.

*Speech.* Speech production skills were evaluated using, among other measures, the conversational speech sample cited above and the *Goldman-Fristoe Test of Articulation-2: Sounds in Words* subtest (*GFTA-2*; Goldman & Fristoe, 2000). As shown in Table 2, several alternative metrics from PEPPER (Shriberg et al., 2001) were used to describe participants' severity of speech impairment, including the Percentage of Consonants Correct (PCC), Percentage of Consonants Correct-Revised (PCCR), Percentage of Vowels Correct (PVC), Percentage of Vowels Correct-Revised (PVCR) and the Intelligibility Index (II). The PCC and PVC score distortions as errors in speech-sound acquisition, whereas the PCCR and PVCR score such responses as correct (viewing the distortions as allophonic detail). Intelligibility Index scores reflect contributions from several domains of verbal development, including the individual and interactive consequences on intelligibility of errors in speech, prosody, voice, and language.

Statistical analyses (two-sample t-tests with unpooled variances) of the data for each of the five speech metrics shown in Table 2 indicated that the speakers with SD had significantly lower scores ( $p < .001$ ) than speakers with TS on all five metrics. For example, the SD speakers'

PCC scores averaged over 20 percentage points lower than those of the TS speakers, classifying them in the mild-moderate to moderate-severe levels on this metric. All consonant, vowel/diphthong, and intelligibility findings for participants with SD were consistent with average values reported for sociodemographically comparable cohorts of children with speech delay of unknown origin (Shriberg & Kwiatkowski, 1994).

*Nonword tasks.* An examiner individually administered the NRT and SRT tasks on the same day the hearing screening was completed. The recorded stimuli for each task were presented free field using a laptop computer and generic tabletop speakers monitored at comfortable loudness levels. The NRT was always administered first, excepting a few occasions when children refused to repeat the NRT stimuli, but agreed to do so after completing several of the other protocol tasks including the SRT. Instructions for each task were presented live voice by the examiner. NRT instructions were as follows: “You will hear a woman’s voice saying some *pretend* words. Say exactly what she says. If you need anything repeated, let me know.” The NRT instructions were also repeated on the laptop presentation by an adult female speaker (fourth author), using a digitized version of the stimuli from the standard (audiocassette) version of the NRT. The SRT instructions presented live by the examiner were as follows: “You are going to say some silly words for me now. Every time you hear the woman say a word, you try to copy her. Say the words exactly the way she says it.” These instructions were not repeated by the adult female speaker who said the 18 SRT stimuli on the laptop presentation. For both tasks, replays of stimuli or requests for replays were provided only when the child clearly was inattentive, there was talker overlap, or when other noise interfered with the stimuli. Thus, excepting these conditions and no matter how ill-formed the child’s attempt to repeat the stimulus, participants were generally permitted only one presentation of each nonsense word.

Both tasks were transcribed by two experienced transcribers using the rules developed for research in child phonology (Shriberg et al., 2001). The transcription and scoring procedures were consistent with those conventionally used to transcribe and score NRT responses, which ignore speech sound distortions and additions, scoring only speech sound deletions and speech sound substitutions as errors (see the SRT Technical Report; Shriberg & Lohmeier, 2008). Thus, scoring procedures were consistent with the first of the four approaches to scoring nonword tasks discussed previously. An interjudge agreement estimate (88%) reported in Shriberg & Lohmeier (2008) supports the reliability of the phonetic transcriptions used to score all nonword task responses.

*Language.* Screening of receptive language skills using the *Clinical Evaluation of Language Fundamentals-Preschool (CELF-P;* Wiig, Secord, & Semel, 1992) had been completed to ensure that for the focus of the physiology study, all participants scored within normal limits (at or above a standard score of 85) for receptive language. Expressive language skills were analyzed using the first 50 complete and intelligible utterances from a 15-minute conversational speech sample (*Systematic Analysis of Language Transcripts [SALT];* Miller & Chapman, 2004). The two SALT variables used to quantify expressive language production status were mean length of utterance in words ( $MLU_w$ ) and lexical diversity (LD: number of different words; see Klee, Gavin, & Stokes, 2007; Klee, Stokes, Wong, Fletcher, & Gavin, 2004).

As shown in Table 2, children in the four SD age groups had lower average scores on each  $MLU_w$  and LD comparison than participants in the TS groups and were significantly lower at the “Total” group (i.e., across all four age groups) level ( $MLU_w$ :  $t = 4.05$ ;  $p = .001$ ; LD:  $t = 5.68$ ;  $p = .001$ ). It is useful to note here that average  $MLU_w$  scores for the TS speakers were

notably lower than those included as SALT reference data (Miller & Chapman, 2004). Differences in elicitation procedures and SALT coding procedures were ruled out as possible sources for these findings, as methods were similar. Rather, likely interdependent sources for the lowered MLU in word values in the present data are the relatively small sample size of children in the Madison, Wisconsin area database (Miller & Chapman, 2004) compared to the present sample size and sociodemographic considerations. The average MLUw scores from the TS speakers in Table 2 are consistent with values from a considerably larger ( $N > 700$ ) demographically representative sample of children from the Pittsburgh, Pennsylvania area (Dollaghan, 2004; Paradise et al., 2001; 2003;). Although these comparison data reported MLU in morphemes rather than words, MLUw data for this sample were available to the fourth author (CD). The average MLUw scores and standard deviations from this large sample at age 36 months ( $N=748$ ) and at age 48 months ( $N=731$ ) were 2.61 (0.62) and 3.15 (0.66), respectively; the average MLUw scores from the present TS speakers are consistent with these values.

#### *Summary of Psychometric Findings*

The SRT Technical Report (Shriberg & Lohmeier, 2008; Tables 1 and 2) includes descriptive and inferential statistics for the distributions of total NRT and SRT scores and scores at each level of syllabic difficulty. Table 3 is a summary of these comparative analyses, and other psychometric findings. As detailed in the Technical Report and summarily concluded in Table 3, the distributional statistics, internal reliability, and concurrent validity findings for the SRT are interpreted as providing psychometric support for its use in parametric statistical designs. Overall, findings support use of the SRT as a nonword repetition task with speakers as young as three years of age who have either typical or delayed speech sound acquisition.

## Results

*Question 1. Does the SRT eliminate the speech confound when assessing speakers with misarticulations?*

### *Goal and Method*

The goal of the first construct validity study of the SRT, a proof of concept study, was to assess the validity of the claim that there is a significant speech confound in nonword repetition testing with young children and children with speech delay. We addressed this question by asking whether children with typical speech or speech delay had in their phonetic inventories the five phonemes used in the SRT and the 20 phonemes used in the NRT. Due to the extensive data reduction needs, a subsample of 99 (62.7 %) of the 158 participant records was examined. These 99 records were from the first 100 participants (data on one participant was later lost to analysis due to a classification error) on whom data analyses were completed in the original physiology study. Statistical analyses of the eligible comparisons yielded no significant between-group (i.e., subsample compared to remaining participants) mean differences in scores on the two language and five speech metrics described in the Assessment section.

To provide the most liberal test of the hypothesis (i.e., to minimize the likelihood of a Type I error), a speech sound was considered to be ‘in’ a child’s phonetic inventory if it was transcribed as correct at least one time in any one of several speech tasks in the assessment protocol. Thus, to quantify the magnitude of the potential speech confound we first tallied the number of children in the subsample who articulated each of the sounds included in the SRT and NRT correctly at least once in their GFTA-2 responses. If not correct or targeted in the GFTA-2, we then inspected each participant’s conversational speech sample using the same criteria. If not

correct or targeted in conversation, we then inspected transcriptions of each of the several supplementary speech tasks included in the protocol for the physiology study.

### *Results*

*SRT.* Analyses of the 99 subsample participants indicated that /e/ and each of the four consonants used in the SRT stimuli were observed to be articulated correctly in the conversational samples of each participant and hence was considered to be in the phonetic inventories of all of the 99 participants.

*NRT.* Table 4 is a summary of the phonetic inventory findings for the NRT. The left-most column lists those phonemes that were not in the phonetic inventory of at least one participant. The next group of columns provides information for the frequency of occurrence of each target phoneme in the stimuli within each syllable level (i.e., 1- 4 syllables; total targets) in the NRT. The right-most columns include the percentage of participants (by speaker and age groups arranged in decreasing order) who did not have the phoneme attested as correct at least once in any of the several speech sources examined.

Beginning with the diphthong /ɔɪ/, which occurs at each syllable level in the NRT for a total of 8 occurrences (see appendix), 6% to 46% of participants in 5 of the 8 subgroups did not have this sound attested as in their phonetic inventories. The posterior stop /k/, which occurs once in a 2-syllable NRT item, was not in the phonetic inventories of 6% and 8% of children in two subgroups. The labiodental fricatives /f/ and /v/, which occur in two to all four NRT syllable levels in a total of 13 of the 16 NRT stimuli, were not attested as in the inventories of 11% and 31% of children in two subgroups. Finally, the affricates /tʃ/ and /dʒ/, which occur in one to three syllable levels in 10 of the 16 NRT stimuli, were not attested as in the phonetic inventories of 6% to 38% of children in 5 of the 8 subgroups.

### *Conclusion*

The phonetic inventory findings for children with SD in Table 4 are interpreted as support for a speech confound in nonword repetition scores used to date in genetic analyses and in other research contexts. Across the four SD age groups, 6 of the phonemes occurring in the NRT stimuli could not be attested (using the minimum requirement of correct at least once) as ‘in’ the phonetic inventory of all participants, including children in the oldest of the four subgroups. Thus, as shown in the last column of Table 4, from 6% to 46% of the participants’ incorrect repetitions (i.e., deletions of or substitutions for), these sounds in nonwords were likely due to misarticulations, rather than inability to accurately repeat these sounds as presented in nonwords. Much as typical speakers would not be expected to repeat correctly speech sounds that are not phonemic in their native language, such errors cannot be attributed to one or more pre-execution speech processing constraints. A generalization from these findings is that such scoring and interpretive confounds are proportionally more frequent and consequential in nonword repetition tasks that include additional consonants that children often misarticulate in singleton and/or cluster contexts. Moreover, this confound might be expected to be more pronounced in speakers with SD who have significant and difficult-to-transcribe errors on vowels and diphthongs, such as is commonly reported for speakers with cognitive disability, hearing loss, and motor speech disorders.

*Question 2. Does the SRT identify children with expressive language impairment?*

### *Goal and Method*

The second validity study of the SRT addressed the increasing use of nonword repetition tasks for clinical purposes, as reviewed previously — to identify speakers who have, or who are at risk for, language impairment. Two methodological limitations in the present estimate of the

SRT's accuracy in identifying expressive language impairment are important to underscore. First, due to the exclusion of participants with receptive language impairment in the physiology study from which the present data were obtained, this estimate of the sensitivity of the SRT to identify language disorder is limited to disorder in language expression. Second, the only two domains in the corpus used to classify expressive language disorder were  $MLU_w$  and lexical diversity (i.e., Number of Different Words) as quantified using SALT analyses procedures. As described previously, we classified children as having Expressive Language Impairment when their  $MLU_w$  and/or NDW values were more than one standard deviation below the mean from their TS age group counterparts.

### *Results*

The upper section of Table 5 is a summary of descriptive and inferential statistical findings assessing each nonword task's ability to discriminate children with Typical Language (TL) from those with Expressive Language Impairment (ELI) as defined for this study. Between-group differences were tested for statistical significance and characterized in magnitude using three sets of statistics: effect size (Cohen's  $d$ ) for group mean differences, positive and negative likelihood ratios, and diagnostic accuracy percentage (i.e., the percentage of participants classified correctly, including Fisher Exact tests of significance). For each of the 10 between-group mean comparisons in the upper section of Table 5, participants with ELI scored lower than those with TL on both nonword repetition tasks. Except for the comparison between SRT scores for participants in the oldest age group, for which there were only 6 participants with ELI, all statistical tests for these differences were significant and effect sizes were primarily *large*, with across age group totals of 0.88 for the NRT and 0.70 for the SRT. Recall previously, however,

that the average effect size was 1.30 for this comparison in the 23 studies included in the Graf Estes et al. (2007) meta-analysis.

Positive and negative likelihood ratios and diagnostic accuracy percentages for each nonword repetition task were computed by a software utility that obtained these statistics using logistic regression. For each task, the utility determined the point in the distribution of scores at which the difference between two fractions yielded the smallest absolute value: fraction present below the score =  $\text{False Negatives}/(\text{True Positives} + \text{False Negatives})$  and fraction absent above the score =  $\text{False Positives}/(\text{False Positives} + \text{True Negatives})$ . The resulting positive likelihood ratios for the SRT and NRT ranged from 1.0 (indicating a lack of diagnostic value for identifying children with ELI) to 4.64, a value that is considered moderately positive, with intermediate diagnostic impact (Sackett et al., 1991; Straus, Richardson, Glasziou, & Haynes, 2005). Similarly, the negative likelihood ratios for the two tasks ranged from 1.0 (indicating a lack of diagnostic value for identifying children with typical speech) to 0.25, a moderately negative value that is also classified as having intermediate impact. It should be noted that confidence intervals for all positive and negative likelihood ratios of the SRT and NRT overlapped at every age level, suggesting that the two tasks were of comparable diagnostic value with respect to ELI. Notably, the two nonword tasks had essentially similar diagnostic accuracy. Over all age groups, as indicated in the right-most two columns of the upper section in Table 5, each of the two tasks was approximately 68% accurate in discriminating participants with expressive language impairment.

The other two sections in Table 5 provide additional descriptive and inferential statistics, with participants aggregated by Typical Speech versus Speech Delay (middle section) and Typical Language-Typical Speech versus Typical Language-Speech Delay (lower section).

These analyses assess whether the nonword tasks discriminated participants with speech delay, with or without ELI. Findings were generally comparable for the two nonword repetition tasks. On each of the between-group comparisons, participants with speech delay, with and without language impairment, had significantly lower nonword repetition scores than participants with typical speech.

### *Conclusion*

In the present study, the obtained diagnostic accuracy value for the SRT of 68%, was comparable to the accuracy value obtained for the NRT. We view this interim finding as a minimal estimate of the expected increased sensitivity/specificity of the SRT to identify language impairment because this task is not confounded by the speech production errors that lower NRT scores yielding potentially false positive classification for language impairment in speakers with misarticulations (see findings for Question 1). As noted previously, the lack of participants with receptive language impairment in the present study and the lack of information on additional variables to classify participants' expressive language status likely attenuated both the between-group differences in scores on the two nonword repetition tasks and the effect sizes and diagnostic accuracy findings associated with those differences. The validity of this constraint on the language classifications in the present paper is supported by average between-group difference in NRT scores for participants with typical and impaired language in the present study compared to the NRT findings reported in Ellis Weismer et al (2000). As reviewed previously, average NRT scores for children with language impairment as defined in that study were nearly three standard deviations below those obtained for children with typical language. In the present study, as just reviewed, the difference in magnitude over all four age groups was less than one standard deviation for both the SRT and the NRT (0.88 and 0.79, respectively).

*Question 3. Does the SRT provide information on speech processing constraints that may underlie poor nonword repetition task performance?*

The goal of the third study was to assess the informativeness of the SRT in providing data and insights on alternative speech processing constraints underlying poor nonword repetition performance. Analyses of all participant data indicated that 88.4% of the consonant repetition errors were speech sound substitutions, 4.8% were speech sound deletions, and the remaining 6.8% were non-attempts. These findings were viewed as support for the potential utility of item analyses of the substitutions errors (Marton & Schwartz, 2003) within each of the putative speech processing domains. It is efficient to present findings for speech processing in the reverse order in which they are typically schematized, beginning with findings for prearticulatory planning/programming processes.

*Question 3a. Do findings from an item analyses of the SRT error data support a planning-programming constraint in nonword repetition?*

*Goal and method.* One potential source of the lowered SRT scores obtained for some of the present participants might occur at the processing phase immediately preceding overt speech – the pre-execution level of planning gestures and programming movements (van Lieshout, Bose, Square, & Steele, 2007; van Lieshout & Goldstein, 2008). The *ease of articulation* construct (e.g., Locke, 1972) would predict that it would be more difficult to plan/program the movements for two different consonants that differ in both place (heterorganic) and manner (heterotypic) than to plan/program movements for two different consonants that share one feature (i.e., are homorganic or homotypic). The four consonants included in the SRT, two voiced stops and two nasals, allow a test of this perspective as described below.

*Results and conclusion.* Table 6 includes item-level nonword repetition performance findings for the eight 2-syllable SRT stimuli (to minimize contributions from perceptual and memorial constraints associated with the 3- and 4-syllable stimuli) listed in the order in which they were presented (see appendix). Notice that the two target consonants in the three bolded stimuli are both heterorganic and heterotypic (deme, mede, and nebe). Of the remaining stimuli, two include consonants that differ only in place (bede, debe) and three include consonants that differ only in manner (beme, nede, and mebe). Table 6 includes mean percent correct scores on each item individually for participants in the four subgroups and overall. After each mean performance, the rank ordering for each item (“1” = highest performance) is listed in square brackets.

As assessed using the ease of articulation-based metric and the limited number and complexity of responses in the 18 SRT items, the nonword repetition performance findings in Table 6 do not support a planning/programming constraint underlying the repetition errors in SRT responses. In the rightmost “All” column, performance on the nonwords posited to be most difficult to articulate because their consonants differed in both place and manner (deme, mede, and nebe) were ranked 1<sup>st</sup>, 4<sup>th</sup>, and 5<sup>th</sup> in accuracy, respectively, among the 8 stimuli. There were no notable differences in these overall rankings within the four subgroups. Thus, there was no apparent association between the number of feature differences in target consonants and their repetition accuracy. Again, the present stimuli were, limited to four early-occurring sounds in singleton contexts, possibly failing to reach the level of phonetic complexity needed to challenge articulatory planning/programming processes.

*Question 3b. Do findings from an error analysis of the SRT data support a memory capacity constraint in nonword repetition?*

*Goal and method.* As reviewed previously, the predominant perspective in the literature is that a memory capacity limitation affecting storage and/or retrieval processes underlies poor nonword repetition task performance (Gathercole & Baddeley, 1990). Support for such memorial constraints in the present data can be marshaled from several of the present findings.

*Results and conclusion.* First, as invariably reported in the literature, the average nonword repetition scores of participants responding to both tasks were lower as nonword stimuli increased in length (Table 3). Additional analyses indicated that only 2.6% of participants had higher repetition accuracy on the 3-syllable nonwords than on the 2-syllable nonwords, and only 16.0% had higher accuracy on the 4-syllable nonwords than on the 3-syllable nonwords. Because these percentages reflecting counter support for a memorial constraint are not inordinately large (with some accuracy reversals plausibly due to external factors such as momentary inattention), the stimulus length (i.e., difficulty level) findings summarized in Table 3, and presented in detail in the Technical Report, are interpreted as support for memory capacity limitation as, at least in part, a processing constraint underlying nonword repetition errors.

The appropriate interpretation of data from another finding that addresses the memory constraint hypothesis, however, is less clear. Analyses indicated that 38 of the 156 (24.4%) participants had scores ranging from 0% to 50% correct on the 2-syllable SRT stimuli. Thus, a nontrivial percentage of children in this study (approximately one-quarter of the sample) had difficulty in correctly repeating nonwords that required storage and retrieval of only two consonant singletons that, as shown in the first study, they did not misarticulate. The majority of these participants (33 of the 38 children, 87%) had SD, including 16 (42.7%) with SD and typical language and 17 (44.7%) with SD and expressive language impairment.

The latter findings above are viewed as counter support for the interpretation of nonword repetition tasks as assessing only phonological working memory. Repetition of simple CVCV stimuli using one constant vowel in all stimuli would appear to make minimal demands on memory capacity, implying that poor performance on such stimuli is associated with one or more alternative processing constraints. Graf Estes et al. (2007) have reviewed similar counter support for memory capacity based on repetition errors reported for stimuli as short as 1-syllable nonwords in many studies. Thus, especially in consideration of the findings discussed next, results from these analyses add to literature trends reviewed previously indicating that nonword repetition task performance is multiply determined.

*Question 3c. Do findings from an error analysis of the SRT data support an auditory-perceptual encoding constraint in nonword repetition?*

*Goal and method.* A third analysis addressed the hypothesis reviewed previously of an auditory-perceptual processing constraint affecting the veridical encoding of short-term representations of nonwords. The presence of such a constraint might be identifiable by inspecting the manner feature of all substitution errors in nonword repetition responses. Preservation of the manner feature of the target consonant (i.e., a *within-class* repetition error) might be interpreted as indicating at least partial encoding of the target consonant, whereas *between-class* repetition errors might be interpreted as evidence of a constraint in the auditory-perceptual component of encoding.

To assess the level of statistical support for this possible processing constraint, all substitution errors on the SRT were divided into two classes: within-class errors (i.e., stops for stops; nasals for nasals) and between-class errors (i.e., nasals or some other manner class for stops; stops or some other manner class for nasals). Of the 50 consonant targets in the SRT, 28

were stops (b, d) and 22 were nasals (m, n). The percentages of within-class and between-class errors for stops and nasals were determined individually for each participant in each of the four speech-language groups (TL-TS, TL-SD, ELI-TS, ELI-SD) by dividing a participant's total numbers of substitution errors of each type by his or her total number of substitution errors. Preliminary inspection of these data indicated that the substitution error data for 2-syllable words were too sparse for parametric statistical analyses, but the error data were adequate for analyses at the 3-syllable, 4-syllable, and the Total of all levels combined. To assess the null hypothesis of no between-group differences in the percentage of within-class substitution errors among the four speaker groups, three general linear ANCOVA models compared their error types on the 3-syllable SRT stimuli, the 4-syllable stimuli, and all SRT stimuli, with gender and age as covariants.

*Results and conclusion.* Figure 1 includes the gender and age-adjusted means and 95% confidence interval bars around the means for the within-class substitution errors of speakers in the four language-speech status groups. The underbars indicate all significant between-group effect sizes adjusted for multiple comparisons (Hochberg & Tamhane, 1987) and their confidence boundaries. In comparison to the cell sizes in the other three groups (ranging from 37-53 participants), there were relatively few speakers in the group with ELI and TS (9-10 participants).

The statistical findings and descriptive trends in Figure 1 are interpreted as preliminary support for an auditory-perceptual encoding constraint as a possible source of repetition errors in participant's with typical language and speech delay (TL-SD) and expressive language impairment and speech delay (ELI-SD), as defined in this study. As assessed in 3-syllable, 4-syllable, and Total SRT stimuli, the within-class substitution errors of participants with typical

language and typical speech (TL-TS) averaged 59.5%, 67.2%, and 61.6%, respectively. In comparison, within-class substitution errors at these difficulty levels for participants with typical language and speech delay (TL-SD) averaged 43.9%, 55.2%, and 51% respectively, and speakers with expressive language impairment and speech delay (ELI-SD) averaged 45.4%, 43%, and 44.1%. All six effect size comparisons (ranging from .47 to .90) were significant, with higher effect sizes more associated with two of the three TL-TS and ELI-SD comparisons. Only one of the three effect sizes comparing the participants with TL-TS to participants with expressive language impairment and typical speech (ELI-TS) was significant (all SRT words: ES = .73). As above, effect sizes for all comparisons with this group were likely attenuated due to small cell sizes.

The data shown in Figure 1 appear to be the first nonword item analyses of featural data supporting a possible auditory-perceptual processing constraint in children with speech-language impairment. Among alternative interpretations for these findings, participants with higher within-class error rates may have correctly perceived and encoded the correct manner features for consonant targets, but this information was subsequently lost due to memory constraints. Such an interpretation would need to account for the present findings (and those noted previously for 1- and 2-syllable stimuli) indicating that substitution errors were not limited to the longer SRT stimuli—they were observed on the presumably minimally memory taxing 2-syllable SRT stimuli.

### Summary and Research Needs

#### *Summary*

This report has summarized information on the construction and psychometric properties of a nonword repetition task and presented findings from three validity studies addressing its use

in genetic and other research in which participants have misarticulations or can be assessed more readily with minimally-demanding speech stimuli. Findings summarized from the psychometric analyses reported in the SRT Technical Report support the validity and reliability of the SRT as an endophenotype for use in parametric statistical designs in genetic and other research. For studies focused on speech rather than language, it is useful to note that analyses described in the Technical Report indicate that speakers with SD in this study had significantly lower SRT scores (62.3% ) compared to speakers with TD (75.6%;  $t = 4.59$ ;  $p = .001$ ; effect size [Hedges corrected] =  $-.75$ ; C.I.:  $-0.42/-1.08$ ). Analyses described in the Technical Report also support the administrative efficiency of the SRT. Briefly, (a) there were relatively few non-responses to SRT items, (b) all but two participants completed the task, (c) administration times were less than two minutes for children in all age groups, and (d) transcription and scoring times were brief and efficient using both manual procedures and software utilities.

Findings from three studies are interpreted as support for the construct validity of the SRT. Results from a proof of concept study indicated that whereas young speakers with typical and delayed speech had the five phonemes used in the SRT in their phonetic inventories, a substantial proportion of participants with speech delay did not have in their phonetic inventories all of the phonemes used in a comparison nonword repetition task. Results from a comparative diagnostic accuracy study indicated that both the SRT and the NRT were approximately 68% accurate in identifying young children with expressive language impairment as defined for the participants this study. Results from a third study series provided preliminary information on possible sources of processing constraints underlying substitution errors on consonant sounds in children with speech delay and/or expressive language impairment. Within the methodological caveats described for each analysis and to be discussed next, findings were interpreted to provide

no support for a constraint in planning the gestures and programming the movements for speech sound production, mixed support for the sufficiency of a constraint in storage and/or retrieval processes, and interim support pending cross-validation for an auditory-perceptual encoding constraint contributing to nonword repetition errors in children with speech delay and expressive language impairment.

### *Research Needs*

The primary motivation for the development of the SRT was to address a significant methodological confound in speech-genetics research. A brief review of interpretative issues in this literature when reporting nonword repetition task findings illustrates some of the consequences of this constraint. Stein et al. (2004) reported that for a battery of verbal tasks administered to children with SD, the most significant linkage findings with several chromosome 3 susceptibility loci were obtained for the nonword repetition task developed by Kamhi and Catts (1986). Owing to the procedures used to score responses on this repetition task, which considered any deletion or substitution of a sound in a response as incorrect, the authors included the following caveat in a discussion of the linkage findings: “Since our variables for phonological memory . . . are correlated with the articulation measures, we are unable to specify whether the linkage with the [articulation] factor is due to articulation alone” (p. 295). Similarly, in a study of 4.5-year-old twins, Kovas et al. (2005) noted:

Because children’s articulation at this age is often immature, it was not feasible to adjust scoring to allow for misrepetitions that were consistent with the child’s expressive phonological repertoire. Thus, results from this measure [a nonword repetition task] are sensitive to articulatory accuracy as well as phonological short-term memory. (p. 638-639)

The SRT may provide the means to disambiguate the contribution of articulation errors in nonword repetition tasks from those of one or more speech processing sources of poor performance. The present findings are viewed as a conservative estimate of the magnitude of such constraints because one of the goals in development of the comparison measure, the NRT, was to minimize the articulatory demands evident in other nonword repetition measures. Thus, the phonetic inventory findings reported for the first study comparison presumably would be more substantial if SRT findings were compared to phonetic inventory findings from each of the other nonword repetition tasks reviewed in Graf Estes et al. (2007) and others used to date in speech-genetics studies. However, it is appropriate to note several methodological and substantive research needs that limit generalizations from the present findings.

*Methodological needs.* One methodological need is for a study addressing possible order effects in the present data. The fixed order in the present protocol, with all participants administered the NRT before the SRT, could have influenced performance on either or both measures. Second, it is clear that the restricted stimulus set in the SRT creates a dense phonological neighborhood for participants in which correct responses require inhibition of the memory trace from each of the preceding syllables/items. Such potential effects can be explored using alternative ordering of the 16 stimuli. A third need is to obtain estimates of the test-retest stability of SRT scores, especially as they may be used to provide estimates of the standard errors of measurement by age and difficulty level. Gray (2003) reports that nonword performance was significantly improved on a second administration, but remained relatively stable on a third administration. A fourth need, especially for additional study of auditory-perceptual issues in speech processing, is to obtain acoustic descriptions of each of the SRT consonant stimuli spoken by the adult female. Findings reported in Edwards, Beckman, and

Munson (2004, Table 3) illustrate the possible influence of durational differences in phonemes used in nonword repetition tasks on performance, with the nasals /m/ and /n/ grouped among what the authors termed the “short” segments. A fifth methodological need is to increase the sensitivity of the SRT for use with older respondents with typical speech, possibly by adding 5-syllable nonword items. As reported for other nonword tasks, ceiling effects for typical speakers were obtained in the developmental comparison data reported in Potter et al. (2008) and included in the SRT Technical Report (Shriberg & Lohmeier, 2008). On this issue, it is important to underscore that in comparison to the available developmental data for the SRT and NRT, the NRT has higher sensitivity to individual differences in speech processing in older, typical speakers (Potter et al., 2008).

*Substantive needs.* Elsewhere we have reported preliminary structural equation modeling of the present data to attempt to understand the role of expressive language impairment as a mediating variable in the association between familial risk for speech sound disorder and lowered nonword performance in children with typical speech and speech delay (Shriberg, Lohmeier, Dollaghan, Campbell, & Moore, 2007). Findings indicated that expressive language impairment was a significant mediator of lowered nonword repetition performance in both speech groups, but again, findings were constrained by the sensitivity of the language measures available in these data. A major substantive need is to cross-validate such findings and the estimates of the SRT's ability to discriminate language impairment in replication studies in which participants' receptive and expressive language impairment are well characterized. As noted by Coady and Evans (2008), particularly in their discussion of findings by Briscoe et al. (2001), significant statistical associations between nonword repetition ability and vocabulary appear to hold for measures of receptive vocabulary (cf. Bowey, 1996, 1997, 2001), but not for measures

of expressive vocabulary (however, see Lahey & Edwards, 1997). Due to the increased risk for misarticulations in young children ascertained for language disorder using clinical samples (Bishop & Hayiou-Thomas, 2008; Tomblin et al., 1997), such questions about nonword performance and language domains have not been assessed without the speech confound that motivated development of the SRT.

Another need is for developmental reference data from speakers with histories of both normalized and persistent speech-language disorder. As noted earlier, a likely reason why nonword repetition tasks have been productive in genetics research is that they appear to be sensitive to speakers who are at risk for any of five types of verbal trait disorders (speech, language, reading, spelling, writing), who have active verbal trait disorders, or who have normalized prior verbal trait disorders. Such endophenotypic sensitivity has several methodological and conceptual advantages over the clinical-behavioral phenotypes used to classify and quantify active disorders. Specifically, the availability of this information for family members in three or more generations contributes significantly to the power of detecting a disorder segregating in a family. Consistent with the concept of 'generalist genes' (Plomin & Kovas, 2005) underlying verbal trait disorders, scores on nonword repetition tasks appear to be sensitive to neurolinguistic processing across the lifespan, while perhaps not specific for any one verbal trait or clinical phenotype.

### Conclusion

Based on SRT findings reported in the present paper and elsewhere (Preston, 2008; Shriberg et al., 2006; Shriberg, Jakielski, & El-Shanti, 2008; Shriberg, Paul, Black, & van Santen, 2009; Shriberg, Potter & Strand, 2009), this task appears to be useful in genetic and other research with children and adult participants who have incomplete phonetic inventories or speech

disorders of known or unknown origin. We would also note here several collaborative studies in progress with participants who speak a number of American, Australian, and Scottish dialects, made possible by the lack of scored vowel targets in the SRT, and the potential use of the SRT in cross-linguistic studies with children learning the many languages in which the consonants /b/, /d/, /m/, and /n/ are included in the phonetic inventory of the language and mastered early (Locke, 1993; Vihman, 1996).

## References

- Adams, A. M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders*, 35, 95-116.
- Alphen, P. van, Bree, E. de, Gerrits, E., Jong, J. de, Wilsenach, C., & Wijnen, F. (2004). Early language development in children with a genetic risk of dyslexia. *Dyslexia*, 10, 265-288.
- Archibald, L. M. D., & Gathercole, S. E. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research*, 49, 970-983.
- Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36, 189-208.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417-423.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 47-90). New York: Academic Press.
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28-61). New York: Cambridge University Press.
- Bishop, D. V. (2002a). The roles of genes in the etiology of specific language impairment. *Journal of Communication Disorders*, 35, 311-328.
- Bishop, D. V. (2002b). Motor immaturity and specific speech and language impairment: Evidence for a common genetic basis. *American Journal of Medical Genetics*, 114, 56-63.

- Bishop, D. V. M., Adams, C. V., & Norbury, C. F. (2004). Using nonword repetition task to distinguish genetic and environmental influences on early literacy development. *American Journal of Medical Genetics, 129B*, 94-96.
- Bishop, D. V. M., Bishop, S. J., Bright, P., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language, and Hearing Research, 42*, 155-168.
- Bishop, D. V. M., & Edmundson, A. (1987). Specific language impairment as a maturational lag: Evidence from longitudinal data on language and motor development. *Developmental Medicine and Child Neurology, 29*, 442-459.
- 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.
- Bock, J. K. (1982). Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*, 1-49.
- Bonnet, B. A. (1999). *Performance of typically developing children and children with phonological disorders on a nonword repetition task*. Unpublished master's thesis, Ohio State University, Columbus.
- Bowey, J. A. (1996). On the association between phonological memory and receptive vocabulary in five-year-olds. *Journal of Experimental Child Psychology, 63*, 44-78.
- Bowey, J. A. (1997). What does nonword repetition task measure? A reply to Gathercole and Baddeley. *Journal of Experimental Child Psychology, 67*, 295-301.

- Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics*, 22, 441-469.
- Briscoe, J., Bishop, D. V. M., & Norbury, C. F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 42(3), 329-340.
- Butcher, L. M., & Plomin, R. (2008). The nature of nurture: A genomewide association scan for family chaos. *Behavior Genetics*, 38, 361-371.
- Campbell, T., Dollaghan, C., Needleman H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research*, 40, 519-525.
- Caylak, E. (2007). A review of association and linkage studies for genetical analyses of learning disorders. *American Journal of Medical Genetics Part B*, 144B, 923-943.
- Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of nonword repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, 43, 1-40.
- Coady, J. A., Evans, J. L., Mainela-Arnold, E., & Kluender, K. R. (2006). Children with specific language impairments perceive speech most categorically when tokens are natural and meaningful. *Journal of Speech, Language, and Hearing Research*, 50, 41-57.
- Coady, J. A., Kluender, K. R., & Evans, J. L. (2005). Categorical perception of speech by children with specific language impairments. *Journal of Speech, Language, and Hearing Research*, 48, 944-959.

- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283-321.
- Dollaghan, C. A. (2004). Taxometric analyses of specific language impairment in 3- and 4-year-old children. *Journal of Speech, Language, and Hearing Research*, *47*, 464-475.
- Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, *41*, 1136-1146.
- Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, *47*, 421-436.
- Edwards, J., Fox, R. A., & Rogers, C. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and phonetic inventory size. *Journal of Speech, Language, and Hearing Research*, *48*, 61-78.
- Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*, *19*, 279-309.
- Ellis Weismer, S., & Edwards, J. (2006). The role of phonological storage deficits in specific language impairment: A reconsideration. Invited commentary on S.E. Gathercole, Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 556-562.
- Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research*, *43*, 865-878.

- Evans, J. L., Viele, K., Kass, R. E., & Tang, F. (2002). Grammatical morphology and perception of synthetic and natural speech in children with specific language impairments. *Journal of Speech, Language, and Hearing Research, 45*, 494-504.
- Gallon, N., Harris, J., & van der Lely, H. (2007). Non-word repetition: An investigation of phonological complexity in children with Grammatical SLI. *Cognitive Linguistics & Phonetics, 21*, 435-455.
- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513-543.
- Gathercole, S. E., & Adams, A.-M. (1993). Phonological working memory in very young children. *Developmental Psychology, 29*, 770-778.
- Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory & Language, 29*, 336-360.
- Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation (2<sup>nd</sup> Edition)*. Circle Pines, MN: American Guidance Service.
- Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry, 160*, 636-645.
- Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*, 177-195.
- Gray, S. (2003). Diagnostic accuracy and test-retest reliability of nonword repetition task and digit span tasks administered to preschool children with specific language impairment. *Journal of Communication Disorders, 36*, 129-151.

- Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*, 594-621.
- 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*, 149-171.
- Hochberg, Y., & Tamhane, A. C. (1987). *Multiple comparison procedures*. Hoboken, New Jersey: John Wiley & Sons.
- Howard, L. A., Binks, M. G., Moore, A. P., & Playfer, J. R. (2000). The contribution of apraxic speech to working memory deficits in Parkinson's Disease. *Brain and Language*, *74*, 269-288.
- Kamhi, A. G., & Catts, H. W. (1986). Toward an understanding of developmental language and reading disorders. *Journal of Speech and Hearing Disorders*, *51*, 337-347.
- Klee, T., Gavin, W. J., & Stokes, S. F. (2007). Utterance length and lexical diversity in American- and British-English speaking children: What is the evidence for a clinical marker of SLI? In R. Paul (Ed.), *Language disorders from a developmental perspective: Essays in Honor of Robin S. Chapman* (pp. 103-140). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Klee, T., Stokes, S. F., Wong, A. M.-Y., Fletcher, P., & Gavin, W. J. (2004). Utterance length and lexical diversity in cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, *47*, 1396-1410.
- Kovas, Y., Hayiou-Thomas, M. E., Oliver, B., Dale, P. S., Bishop, D. V. M., & Plomin, R. (2005). Genetic influences in different aspects of language development: The etiology of language skills in 4.5 year-old twins. *Child Development*, *76*, 632-651.

- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lewis, B.A., Shriberg, L.D., Freebairn, L.A., Hansen, A.J., Stein, C.M., Taylor, H.G., et al. (2006). The genetic bases of speech sound disorders: evidence from spoken and written language. *Journal of Speech, Language, and Hearing Research, 49*, 1294-1312.
- Locke, J. (1972). Ease of articulation. *Journal of Speech and Hearing Research, 15*, 194-200.
- Locke, J. (1993). *The child's path to spoken language*. Cambridge, MA: Harvard University Press.
- Marton, K. (2006). Do nonword repetition errors in children with specific language impairment reflect a weakness in an unidentified skill specific to nonword repetition or a deficit in simultaneous processing? *Applied Psycholinguistics, 27*, 569-573.
- Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46*, 1138-1153.
- Miller, J. S., & Chapman, R. S. (2004). *Systematic analysis of language transcripts (Version 8)*. Madison, WI: Language Analysis Laboratory, Waisman Center.
- Monaco, A. P. and the SLI Consortium (SLIC) (2007). Multivariate Linkage Analysis of Specific Language Impairment (SLI). *Annals of Human Genetics, 71*, 1-14.
- Montgomery, J. W. (1995). Examination of phonological working memory in specifically language-impaired children. *Applied Psycholinguistics, 16*, 355-378.
- Moore, C. A., Campbell, T. F., Shriberg, L. D., Green, J. R., Venkatesh, L., Vick, J., et al. (2006). *Physiologic and behavioral classification of delayed speech*. Paper presented at the 5<sup>th</sup> International Conference on Speech Motor Control, Nijmegen, the Netherlands.

- Munson, B. (2006). Nonword repetition and levels of abstraction in phonological knowledge. Invited commentary on S.E. Gathercole, Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 577-581.
- Munson, B., Baylis, A. L., Krause, M. O., & Kim, D. (in press). *Representation and access in phonological impairment*. In C. Fougerón (Ed.), *Papers in Laboratory Phonology 10: Variation, Detail, and Representation*. New York: Mouton de Gruyter.
- Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research*, 48, 61-78.
- 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 language impairments. *Journal of Speech, Language, and Hearing Research*, 48, 1033-1047.
- Oetting, J. B., & Cleveland, L. H. (2006). The clinical utility of nonword repetition for children living in the rural south of the US. *Clinical Linguistics & Phonetics*, 20, 553-561.
- Paradise, J. L., Dollaghan, C. A., Campbell, T. F., Feldman, H. M., Bernard, B. S., Colborn, D. K., et al. (2003). Otitis media and tympanostomy tube insertion during the first three years of life: Developmental outcomes at the age of four years. *Pediatrics*, 112, 265-277.
- Paradise, J. L., Feldman, H. M., Campbell, T. F., Dollaghan, C. A., Colborn, D. K., Bernard, B. S., et al. (2001). Early versus late tympanostomy tube placement for persistent otitis media: developmental outcomes at age 3 years. *The New England Journal of Medicine*, 344(16), 1179-1195.

- Plomin, R., & Davis, O. S. P. (2006). Gene-environment interactions and correlations in the development of cognitive abilities and disabilities. In MacCabe, J., O'Daly, O., Murray, R.M., McGuffin, P., & Wright, P. (Eds). *Beyond Nature and Nurture in Psychiatry: Genes, Environment and their Interplay* (pp. 35-45). Oxford: Informa Healthcare Medical Books.
- Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, 131, 592-617.
- Pollack, K. E., & Berni, M. C. (2003). Incidence of non-rhotic vowel errors in children: data from the Memphis Vowel Project. *Clinical Linguistics & Phonetics*, 17, 393-401.
- Pollack, K. E., & Keiser, N. J. (1990). An examination of vowel errors in phonologically disordered children. *Clinical Linguistics & Phonetics*, 4, 161-178.
- Potter, N. L., Shriberg, L. D., Hersh, J., Karlsson, H. K., Lohmeier, H. L., McSweeney, J. L., et al. (2008). *Perceptual and acoustic developmental comparative data for the Madison Speech Assessment Protocol (MSAP)*. (Tech. Rep. No. 15). Phonology Project, Waisman Center, University of Wisconsin-Madison. Report in preparation.
- Preston, J. (2008). *Phonological processing and speech production in preschoolers with speech sound disorders*. Unpublished doctoral dissertation, Syracuse University, New York.
- Reuterskiöld-Wagner, C., Sahlén, B., & Nyman, A. (2005). Non-word repetition and non-word discrimination in Swedish preschool children. *Clinical Linguistics & Phonetics*, 19, 681-699.

- Rodekohr, R. K., & Haynes, W. O. (2001). Differentiating dialect from disorder: A comparison of two processing tasks and a standardized language test. *Journal of Communication Disorders, 34*, 255-272.
- Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-Year-Olds: Evidence from typically developing children. *Journal of Speech, Language, and Hearing Research, 47*, 223-234.
- Rvachew, S., Ohberg, A., Grawburg, M., & Heyding, J. (2003). Phonological awareness and phonemic perception in 4-year-old children with delayed expressive phonology skills. *American Journal of Speech-Language Pathology, 12*, 463-471.
- Sackett, D. L., Haynes, R. B., Guyatt, G. H., & Tugwell, P. (1991). *Clinical epidemiology: A basic science for clinical medicine* (2<sup>nd</sup> ed.). Boston: Little Brown.
- Sahlén, B., Reuterskiöld-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Nonword repetition in children with language impairment—pitfalls and possibilities. *International Journal of Language and Communication Disorders, 34*, 337-352.
- Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research, 36*, 105–140.
- Shriberg, L. (2008). Childhood speech sound disorders: From post-behaviorism to the post-genomic era. In R Paul & P Flipsen, eds. *Speech Sound Disorders in children: essays in honor of Lawrence D. Shriberg*, Plural Publishing, Inc. In preparation.
- Shriberg, L. D., Allen, C. T., McSweeny, J. L., & Wilson, D. L. (2001). PEPPER: Programs to examine phonetic and phonologic evaluation records [Computer software]. Madison, WI: Waisman Center Research Computing Facility, University of Wisconsin–Madison.

- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L. (1997a). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research, 40*, 708–722.
- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L. (1997b). The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research, 40*, 723–740.
- Shriberg, L. D., Ballard, K. J., Tomblin, J. B., Duffy, J. R., Odell, K. H., & Williams, C. A. (2006). Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting *FOXP2*. *Journal of Speech, Language, and Hearing Research, 49*, 500-525.
- Shriberg, L. D., Jakielski, K. J., & El-Shanti, H. (2008). Breakpoint localization using array-CGH in three siblings with an unbalanced 4q;16q translocation and childhood apraxia of speech (CAS). *American Journal of Medical Genetics Part A*.
- Shriberg, L. D., & Kwiatkowski, J. (1980). *Natural Process Analysis: A procedure for phonological analysis of continuous speech samples*. New York: Macmillan.
- Shriberg, L. D., & Kwiatkowski, J. (1983). Computer-assisted Natural Process Analysis (NPA): Recent issues and data. In J. Locke (Ed.), *Assessing and treating phonological disorders: Current approaches. Seminars in Speech and Language, 4*, 389–406.
- Shriberg, L. D., & Kwiatkowski, J. (1985). Continuous speech sampling for phonologic analyses of speech-delayed children. *Journal of Speech and Hearing Disorders, 50*, 323–334.
- Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech and Hearing Research, 37*, 1100–1126.

- Shriberg, L. D., Lewis, B. A., Tomblin, J. B., McSweeny, J. L., Karlsson, H. B., & Scheer, A. R. (2005). Towards diagnostic and phenotype markers for genetically transmitted speech delay. *Journal of Speech, Language, and Hearing Research, 48*, 834-852.
- Shriberg, L.D., Lohmeier, H.L., Dollaghan, C.A., Campbell, T.F., & Moore, C.A. (2007, April). Nonword repetition in speech genetics research: Aggregation, marker, and memory findings using the Syllable Repetition Task (SRT). Poster presented at the Afasic Fourth International Symposium Warwick, United Kingdom.
- Shriberg, L. D., & Lohmeier, H. L. (2008). *The Syllable Repetition Task (SRT)*. (Tech. Rep. No. 14). Phonology Project, Waisman Center, University of Wisconsin-Madison.
- Shriberg, L. D., Paul, R., Black, L. M., & van Santen, J. P. H. (2009). *Do some children with autism spectrum disorder have superior auditory-perceptual encoding processes?* Manuscript in preparation.
- Shriberg, L. D., Potter, N., & Strand, E. (2009). *Childhood apraxia of speech and other motor speech disorders in persons with galactosemia*. Manuscript in preparation.
- SLI Consortium. (2002). A genomewide scan identifies two novel loci involved in specific language impairment. *American Journal of Human Genetics, 70*, 384-398.
- SLI Consortium. (2004). Highly significant linkage to SLI1 locus in an expanded sample of individuals affected by specific language impairment (SLI). *American Journal of Human Genetics, 94*, 1225–1238.
- Smith, S. D., Pennington, B. F., Boada, R., & Shriberg, L. D. (2005). Linkage of speech sound disorder to reading disability loci. *Journal of Child Psychology & Psychiatry, 46*, 1057-1066.

- Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholinguistic framework*. London: Whurr.
- Stark, R. E., & Blackwell, P. B. (1997). Oral volitional movements in children with language impairments. *Child Neuropsychology*, 3, 81-97.
- Stein, C. M., Schick, J. H., Taylor, H. G., Shriberg, L. D., Millard, C., Kundtz-Kluge, A., et al. (2004). Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. *The American Journal of Human Genetics*, 74, 283-297.
- Stoel-Gammon, C., & Herrington, P. B. (1990). Vowel systems of normally developing and phonologically disordered children. *Clinical Linguistics & Phonetics*, 4, 145-160.
- Storkel, H. (2004). The emerging lexicon of children with phonological delays: Phonotactic constraints and probability in acquisition. *Journal of Speech, Language, and Hearing Research*, 47, 1194-1212.
- Straus, S. E., Richardson, W. S., Glasziou, P., & Haynes, R. B. (2005). *Evidence-based medicine* (third edition). Edinburgh: Elsevier.
- Tager-Flusberg, H. (2007, June). *Defining the Language and Communicative Impairments in Autism Spectrum Disorders*. Paper presented at the 27<sup>th</sup> meeting of The Symposium on Resesearch in Child Language Disorders, Madison, WI.
- Thal, D. J., Miller, S., Carlson, J., & Moreno Vega, M. (2005). Nonword repetition and language development in 4-year-old children with and without a history of early language delay. *Journal of Speech, Language, and Hearing Research*, 48, 1481-1495.
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40, 1245-1260.

- Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research*, 39, 1284-1294.
- van der Merwe, A. (in press). A theoretical framework for the characterization of pathological speech sensorimotor control. In M.R. McNeil, ed. *Clinical Management of Sensorimotor Speech Disorders*, 3<sup>rd</sup> Edition, Thieme Medical Publishers, New York.
- van Lieshout, P. H. H. M., Bose, A., Square, P. A., & Steele, C. M. (2007). Speech motor control in fluent and dysfluent speech production of an individual with apraxia of speech and Broca's aphasia. *Clinical Linguistics & Phonetics*, 21, 159-188.
- van Lieshout, P. H. H. M., & Goldstein, L. M. (2008). *Articulatory phonology and speech impairment*. Manuscript in preparation.
- Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Cambridge, MA: Blackwell.
- Washington, J. A., & Craig, H. K. (2004). A language screening protocol for use with young African American children in urban settings. *American Journal of Speech-Language Pathology*, 13, 329-340.
- Wiig E. H., Secord W. A., & Semel E. (1992). *Clinical Evaluation of Language Fundamentals-Preschool*. San Antonio, TX: The Psychological Corporation; 1992.
- Ziegler, W. (2006). The internal structure of phonetic representations: Evidence from neurophonetics. *Stem-, Spraak- en Taalpathology*, 14 (supplement), p. 63.

### Figure Legend

Figure 1. Descriptive findings (means, confidence intervals for means) and significant effect size findings for the percentage of within-class nonword substitution errors of participants in the four speech-language status groups. TL-TS: Typical Language-Typical Speech; TL-SD: Typical Language-Speech Delay; ELI-TS: Expressive Language Impairment-Typical Speech; ELI-SD: Expressive Language Impairment –Speech Delay.

## Appendix

**NONWORD REPETITION TASK (NRT) STIMULI**

- |            |                  |
|------------|------------------|
| 1. ne]b    | 9. .in9]te[b     |
| 2. vo[p    | 10. ne].o[v3]b   |
| 3. te[j    | 11. d9]te[vqb    |
| 4. d9]f    | 12. t3]v9].e]g   |
| 5. t3]vek  | 13. v3]te.e]d9]p |
| 6. .o[vqg  | 14. dqvo[n9].ig  |
| 7. vq.e]p  | 15. ne].9]te[vub |
| 8. n9]te[f | 16. tqve.ine]g   |

**SYLLABLE REPETITION TASK (SRT) STIMULI**

- |         |            |              |
|---------|------------|--------------|
| 1. bede | 7. nede    | 13. benede   |
| 2. deme | 8. mebe    | 14. menebe   |
| 3. beme | 9. bemene  | 15. bemedene |
| 4. mede | 10. debeme | 16. denebeme |
| 5. nebe | 11. medebe | 17. menebede |
| 6. debe | 12. nebede | 18. nedemebe |

A Technical Report and a PowerPoint presentation of the Syllable Repetition Task (SRT) can be downloaded without cost from the Technical Reports section of the Phonology Project website: <http://www.waisman.wisc.edu/phonology/> (Shriberg & Lohmeier, 2008; *The Syllable Repetition Task (SRT)*, Technical Report No. 14, The Phonology Project, Waisman Center, University of Wisconsin-Madison). The technical report provides (a) psychometric data on the SRT, (b) statistical findings from several additional analyses of the SRT, (c) comparison data obtained from 70 typically-speaking children from 4-to-16 years of age, (d) administration instructions, (e) scoring instructions, and (f) a form for manual scoring of the SRT.

### Acknowledgments

We thank Joan Kwiatkowski, Jane McSweeny, and Steven Pittelko for their significant contributions to the development of the SRT. We also thank the following colleagues for their assistance with the design and completion of the physiology study and data analysis for the present study: Roger Brown, Christine Hollar, Heather Leavy Rusiewicz, Christie Tilkens, Sonja Wilson, and David Wilson. This research was supported by grants DC00496 and DC00822 from the National Institute on Deafness and Other Communicative Disorders and by core grants to the Waisman Center from the National Institute of Child Health and Development (HD03352) and the General Clinical Research Center at Children's Hospital of Pittsburgh (M01RR00084).

**Table 1.** Summary of sociodemographic data for participants in the Typical Speaking (TS) and Speech Delay (SD) groups.

Group	<i>n</i>	Age (months)			Sex		Maternal Education <sup>a</sup>		Maternal Ethnicity <sup>b</sup>
		<i>M</i>	<i>SD</i>	Range	%Male	% Female	<i>M</i>	<i>SD</i>	
TS									
1	16	38.5	1.5	36-40	25	75	3.8	0.5	EA=100.0%
2	20	44.6	1.7	42-47	35	65	3.5	0.8	EA=75.0% AA=15.0% AsA=5.0% ME=5.0%
3	19	51.2	1.6	48-53	53	47	3.5	0.9	EA=78.9% AA=15.8% ME=5.3%
4	8	55.5	1.2	54-57	50	50	3.5	1.1	EA=87.5% AA=12.5%
Total	63	46.4	6.1	36-57	40	60	3.6	0.8	EA=84.1% AA=11.1% AsA=1.6% ME=3.2%
SD									
1	30	38.8	1.6	36-41	74	26	3.3	0.8	EA=77.4% AA=9.7% L=3.2% AsA=3.2% ME=6.5%
2	20	44.3	1.9	42-47	65	35	3.2	0.8	EA=60.0% AA=35.0% ME=5.0%
3	29	49.9	1.7	48-53	72	28	3.4	0.6	EA=82.8% AA=6.9% AsA=3.4% ME=6.9%
4	16	56.3	1.9	54-60	69	31	2.9	0.9	EA=68.8% AA=25.0% ME=6.2%
Total	95	46.3	6.5	36-60	71	29	3.2	0.8	EA=74.0% AA=16.7% L=1.0% AsA=2.0% ME=6.3%

<sup>a</sup> Maternal Education Levels: 1 - some high school education; 2 - high school graduate; 3 - some college education; 4 – college graduate.

<sup>b</sup> Ethnicity Codes –AA=African-American; AsA=Asian-American; EA=European-American; L=Latino; NA=Native American; PI=Pacific Islander; ME=Mixed Ethnicity

**Table 2.** Expressive language and speech status of participants with Typical Speech (TS) and Speech Delay (SD) for four age groups. Group 1: 36-41 mos.; Group 2: 42-47 mos.; Group 3: 48-53 mos.; Group 4: 54-60 mos.

			Expressive Language Measures from a Conversational Speech Sample				Speech Measures from a Conversational Speech Sample									
			<u>Total MLU<sub>w</sub></u>		<u>Lexical Diversity</u>		<u>PCC</u>		<u>PCCR</u>		<u>PVC</u>		<u>PVCR</u>		<u>II</u>	
Group	Age (mos.)	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
TS																
1	36-40	16	2.4	0.73	61.5	15.6	86.3	4.7	90.6	3.6	95.9	3.5	98.4	1.7	96.4	3.2
2	42-47	20	3.0	0.93	74.8	13.5	89.3	5.3	92.8	3.8	96.6	3.1	97.7	2.5	97.1	3.2
3	48-53	19	2.9	1.04	71.7	18.9	91.7	4.6	94.9	2.8	97.4	1.8	98.2	1.6	97.8	2.6
4	54-57	8	3.4	1.09	78.0	16.9	94.1	3.7	95.3	2.8	98.1	1.9	98.5	2.0	99.7	0.5
Total	36-57	63	2.9	0.97	70.9	16.9	89.9	5.3	93.2	3.7	96.8	2.8	98.1	2.0	97.7	2.9
SD																
1	36-41	30	1.9	0.55	47.4	13.2	68.1	9.3	72.7	10.6	90.2	6.5	92.2	5.7	78.7	14.5
2	42-47	20	2.1	0.52	51.3	12.2	66.8	12.7	71.3	12.9	89.9	6.3	92.6	5.7	87.2	11.3
3	48-53	29	2.4	0.69	58.9	15.8	69.4	8.6	75.0	8.8	92.3	3.4	94.0	3.0	91.3	6.6
4	54-60	16	2.9	0.95	70.9	14.9	74.5	6.8	79.8	7.0	93.9	4.0	95.8	3.9	90.5	9.1
Total	36-60	95	2.3	0.74	55.6	16.2	69.3	9.8	74.3	10.3	91.4	5.4	93.4	4.8	86.3	12.1

*Note.* MLU<sub>w</sub> = Mean length of utterance in words; PCC = Percent Consonants Correct; PCCR = Percent Consonants Correct-Revised; PVC = Percent Vowels Correct; PVCR = Percent Vowels Correct-Revised; II = Intelligibility Index.

**Table 3.** Summary of psychometric findings for the Syllable Repetition Task (SRT).<sup>a</sup> See the SRT Technical Report cited in the text for complete details.

Variable	Analyses	Findings	Conclusion
Distributional Statistics	SRT and NRT descriptive statistics (means, medians, ranges, standard deviations, skew, and kurtosis) for speaker with TS and SD within each age group; distributional tests for nonnormality	Statistical tests supported the normality of SRT score distributions For each speaker type x age group there was: - close agreement between mean and median SRT scores - orderly associations between SRT means and standard deviations - acceptable SRT skew and kurtosis values - orderly SRT means differences, comparable to NRT means across ages groups and by increasing syllable length	Distributional statistics for SRT scores meet customary psychometric requirements for parametric descriptive and inferential statistical analyses and were comparable across ages in relative but not absolute magnitudes (i.e., percentage of phonemes correctly repeated) to scores obtained for the NRT.
Internal Reliability	Pearson correlation coefficients computed separately for the SRT and NRT tasks to assess the association of scores at each level of difficulty (i.e., 1-syllable [NRT only], 2-syllable, 3-syllable, 4-syllable) with total scores (i.e., part-whole reliability).	Part-whole coefficients for the SRT were moderate to high (0.49 to 0.92) and comparable to those obtained for the NRT (0.37 to 0.91).	The moderate to high positive part-whole coefficients for SRT scores meet customary psychometric requirements for adequate internal reliability.
Concurrent Validity	Pearson correlation coefficients computed between participants' total SRT and NRT scores and their scores at each level of difficulty	Total score SRT-NRT coefficients were 0.73 for the speakers with TS and 0.65 for the speakers with SD. Coefficients for scores at each syllable length ranged from 0.28 to 0.73, with most ranging from 0.40-0.60.	The SRT-NRT coefficients provide concurrent validity support for the SRT (see the SRT Technical Report for discussion of attenuation due to the speech confound).

*Note.* NRT = Nonword Repetition Task; TS = typical speech; SD = speech delay

<sup>a</sup> Tabular data and additional analyses of these and other psychometric variables are reported in Shriberg & Lohmeier (2008). *The Syllable Repetition Task (SRT)*, Technical Report No. 14, The Phonology Project, Waisman Center, University of Wisconsin-Madison [<http://www.waisman.wisc.edu/phonology/>].

**Table 4.** Phonemes occurring in the NRT stimuli that were not in all participants' phonetic inventories.

Phoneme	Number of Occurrence(s) in the 16 NRT stimuli <sup>a</sup>					Participants with phoneme not in their inventory			
	1 syll	2 syll	3 syll	4 syll	Total	Speaker Group	Age Group	<i>n</i>	Percentage
/ɹ/	1	1	3	3	8	SD	1	13	46%
						SD	2	13	23%
						SD	4	7	14%
						TS	4	8	12%
						SD	3	18	6%
/k/		1			1	SD	1	13	8%
						SD	3	18	6%
/f/	1	1			2	SD	1	13	15%
						SD	4	7	14%
						SD	3	18	11%
/v/	1	3	3	4	11	SD	1	13	31%
						SD	3	18	11%
/./		2	3	4	9	SD	3	18	33%
						SD	1	13	15%
						SD	4	7	14%
						SD	2	13	8%
/j/	1				1	SD	1	13	38%
						SD	2	13	23%
						SD	4	7	14%
						TS	2	17	12%
						SD	3	18	6%

*Note.* NRT = Nonword Repetition Task; syll = syllable; SD = speech delay; TS = typical speech

<sup>a</sup> See Appendix for list of NRT stimuli.

**Table 5.** Group t-test and effect size data assessing the NRT and SRT tasks' ability to discriminate (a) speakers with typical language (TL) from speakers with expressive language involvement (ELI), (b) speakers with typical speech (TS) from speakers with Speech Delay (SD), and( c) speakers with TL-TS from speakers with TL-SD.

Measure: Age group	Expressive														
	Typical Language			Language Involvement			Effect Size <sup>a</sup>		Likelihood Ratios				Diagnostic Accuracy		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	ES	CI	Pos.	CI	Neg.	CI	%	CI	
<b>NRT</b>															
1	3;0-3;5	34	50.9	17.1	13	32.7	15.4	<b>1.08</b>	0.40/1.75	2.14	1.17/3.92	.455	.196/1.06	68.1*	52.7/80.5
2	3;6-3;11	18	68.9	12.4	22	50.1	17.5	<b>1.20</b>	0.52/1.87	4.64	1.61/13.3	.273	.123/.602	80.0*	63.9/90.4
3	4;0-4;5	35	66.1	15.4	13	53.6	10.2	<b>0.87</b>	0.21/1.53	2.20	1.20/4.05	.449	.194/1.04	68.8*	53.6/80.9
4	4;6+	18	72.5	11.6	6	57.3	21.7	<b>1.01</b>	0.04/1.98	2.00	.843/4.75	.500	.153/1.64	66.7	44.7/83.6
	ALL	105	62.8	17.0	54	47.5	17.9	<b>0.88</b>	0.54/1.22	2.12	1.52/2.94	.466	.311/.696	67.9*	60.0/75.0
<b>SRT</b>															
1	3;0-3;5	34	65.9	17.1	13	44.0	18.3	<b>1.23</b>	0.53/1.94	2.32	1.29/4.16	.370	.135/1.01	69.6*	54.1/81.8
2	3;6-3;11	18	76.4	13.1	22	62.6	20.4	<b>0.78</b>	0.13/1.43	2.74	1.25/6.00	.330	.147/.739	74.4*	57.6/86.4
3	4;0-4;5	35	74.6	15.5	13	59.7	15.5	<b>0.95</b>	0.28/1.61	2.45	1.38/4.34	.337	.122/.932	70.8*	55.7/82.6
4	4;6+	18	79.3	13.7	6	74.0	19.7	0.34	-0.59/1.27	1.00	.397/2.52	1.00	.403/2.48	50.0	29.6/70.4
	ALL	105	72.9	16.0	54	58.9	20.4	<b>0.79</b>	0.45/1.14	2.08	1.51/2.85	.439	.284/.679	67.5*	59.5/74.6
<b>(b) Typical Speech vs Speech Delay</b>															
Measure: Age group	Expressive														
	Typical Speech			Speech Delay			Effect Size <sup>a</sup>		Likelihood Ratios				Diagnostic Accuracy		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	ES	CI	Pos.	CI	Neg.	CI	%	CI	
<b>NRT</b>															
1	3;0-3;5	16	59.7	14.5	30	39.4	16.0	<b>1.29</b>	0.63/1.95	2.93	1.22/7.04	.356	.189/.670	73.9*	58.6/85.2
2	3;6-3;11	20	66.4	14.7	20	50.5	18.2	<b>0.95</b>	0.30/1.60	4.00	1.62/9.87	.250	.102/.616	80.0*	63.9/90.4
3	4;0-4;5	19	74.9	12.2	29	54.8	11.2	<b>1.71</b>	1.04/2.38	3.60	1.47/8.82	.306	.156/.599	77.1*	62.3/87.5
4	4;6+	8	80.1	6.9	16	63.0	15.9	<b>1.20</b>	0.29/2.12	3.00	.874/10.3	.333	.131/.845	75.0*	52.9/89.4
	ALL	63	69.0	14.7	95	50.4	17.2	<b>1.14</b>	0.80/1.48	2.65	1.73/4.06	.389	.280/.542	72.2*	64.4/78.8
<b>SRT</b>															
1	3;0-3;5	16	70.6	16.1	30	56.0	18.1	<b>0.83</b>	0.19/1.46	2.21	1.03/4.75	.451	.248/.821	68.9*	53.2/81.4
2	3;6-3;11	20	74.3	16.1	20	63.4	19.8	0.60	-0.05/1.24	2.28	1.09/4.76	.451	.223/.912	69.2*	52.3/82.5
3	4;0-4;5	19	77.8	13.9	29	65.9	16.9	<b>0.74</b>	0.14/1.34	2.18	1.08/4.42	.454	.251/.818	68.8*	53.6/80.9
4	4;6+	8	89.0	8.4	16	72.5	14.8	<b>1.21</b>	0.30/2.13	3.00	.874/10.3	.333	.131/.845	75.0*	52.9/89.4
	ALL	63	76.3	15.4	95	63.4	18.2	<b>0.75</b>	0.42/1.08	2.55	1.66/3.92	.427	.312/.584	70.5*	62.6/77.4
<b>(c) Typical Language-Typical Speech vs Typical Language-Speech Delay</b>															
Measure: Age group	Expressive														
	Typical Language- Typical Speech			Typical Language- Speech Delay			Effect Size <sup>a</sup>		Likelihood Ratios				Diagnostic Accuracy		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	ES	CI	Pos.	CI	Neg.	CI	%	CI	
<b>NRT</b>															
1 & 2	3;0-3;11	29	65.1	13.8	23	47.2	17.3	<b>1.14</b>	0.55/1.73	2.68	1.42/5.07	.360	.176/.735	73.1*	58.7/84.0
3 & 4	4;0-4;6+	23	77.5	11.2	30	61.3	12.7	<b>1.33</b>	0.73/1.93	2.81	1.37/5.78	.361	.193/.673	73.6*	59.4/84.3
	ALL	52	70.6	14.1	53	55.2	16.3	<b>1.01</b>	0.60/1.41	2.27	1.45/3.54	.436	.285/.668	69.5*	59.7/77.9
<b>SRT</b>															
1 & 2	3;0-3;11	29	74.2	15.6	23	63.6	15.9	<b>0.66</b>	0.10/1.22	2.10	1.13/3.90	.504	.280/.910	67.3*	52.8/79.3
3 & 4	4;0-4;6+	23	81.0	13.8	30	72.5	14.9	<b>0.58</b>	0.03/1.13	1.62	.908/2.88	.602	.356/1.02	62.3	47.9/74.9
	ALL	52	77.2	15.1	53	68.7	15.9	<b>0.55</b>	0.16/0.94	1.59	1.08/2.34	.601	.409/.884	61.9*	51.9/71.1

<sup>a</sup> Significant effect sizes are bolded.

\* Fisher exact test &lt; .05.

**Table 6.** Syllable Repetition Task (SRT) findings for the 2-syllable nonwords ranked by performance. <sup>a,b</sup>

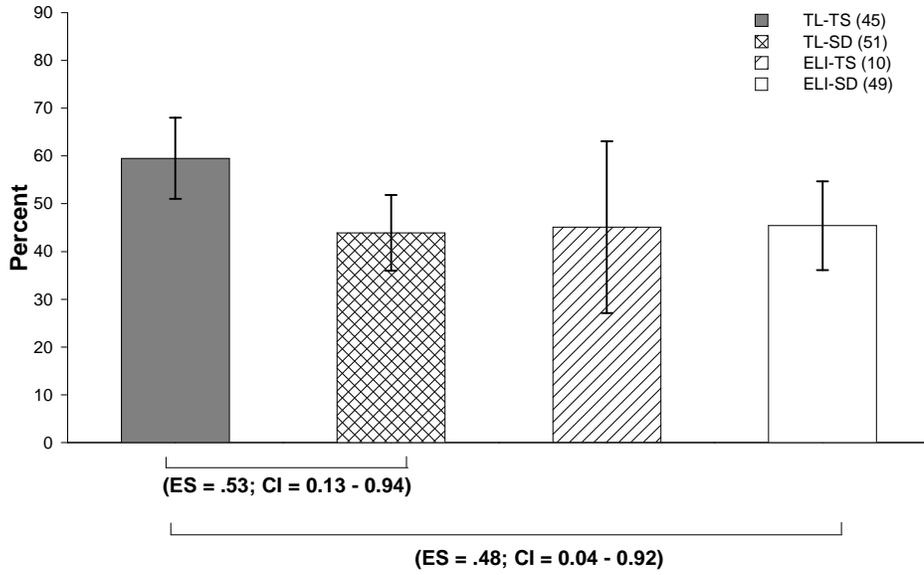
<u>2-syllable SRT nonwords<sup>a</sup></u>	<u>Feature Change</u>		<u>Mean Performance (and rank order in brackets<sup>b</sup>) on each nonword</u>				
	<u>Place</u>	<u>Manner</u>	<u>TL-TS</u>	<u>TL-SD</u>	<u>ELI-TS</u>	<u>ELI-SD</u>	<u>ALL</u>
bede	X		88.5 [4.0]	64.2 [6.0]	90.9 [2.5]	73.2 [2.0]	76.4 [3.0]
<b>deme</b>	X	X	98.1 [1.0]	84.9 [1.0]	90.9 [2.5]	75.6 [1.0]	87.3 [ <b>1.0</b> ]
beme		X	71.2 [8.0]	60.4 [8.5]	90.9 [2.5]	41.5 [7.0]	61.1 [8.0]
<b>mede</b>	X	X	90.4 [3.0]	66.0 [5.0]	72.7 [6.0]	63.4 [4.0]	73.9 [ <b>4.0</b> ]
<b>nebe</b>	X	X	82.7 [6.0]	75.5 [3.0]	72.7 [6.0]	39.0 [8.0]	68.2 [ <b>5.0</b> ]
debe	X		96.2 [2.0]	81.1 [2.0]	90.9 [2.5]	65.9 [3.0]	82.8 [2.0]
nede		X	80.8 [7.0]	71.7 [4.0]	63.6 [8.0]	43.9 [6.0]	66.9 [6.5]
mebe		X	84.6 [5.0]	60.4 [8.5]	72.7 [6.0]	51.2 [5.0]	66.9 [6.5]

*Note.* TL-TS = Typical Language-Typical Speech; TL-SD = Typical Language-Speech Delay; ELI-TS = Expressive Language Impairment-Typical Speech; ELI-SD = Expressive Language Impairment-Speech Delay; ALL = All participants

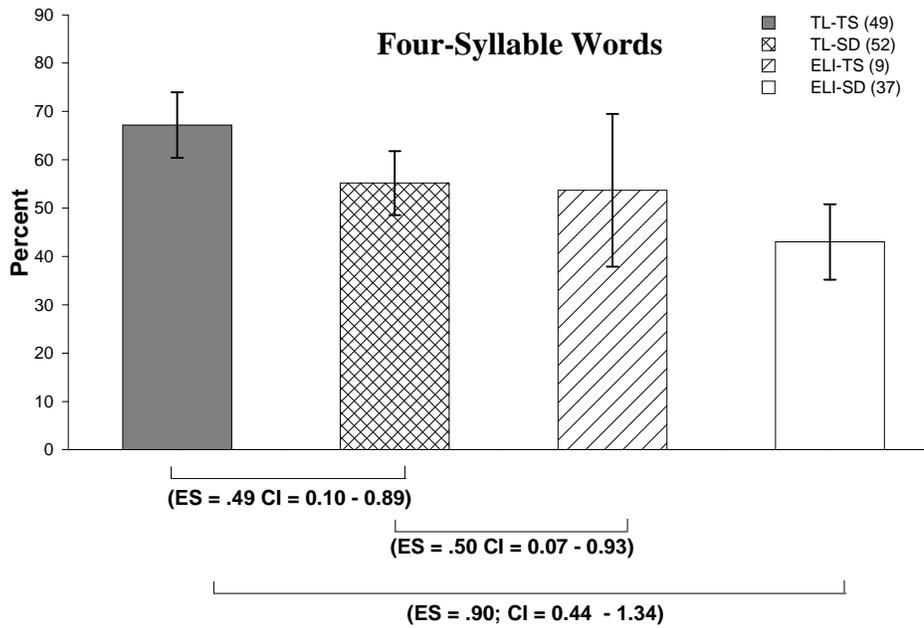
<sup>a</sup> Bolded = 2-feature change

<sup>b</sup> "1" = highest performance

### Three-Syllable Words



### Four-Syllable Words



### All Words

