

Phonemic Awareness Assessment of Children With and Without Speech Sound Disorders: A Comparison Study

Donna Thomas¹, Dee M. Lance², Jacqueline L. Rainey³

¹Department of Health and Human Sciences, Southeastern Louisiana University, Hammond, LA; ²Department Chair of CSD/Speech Language Hearing Center, University of Central Arkansas, Conway, AR; ³College of Health and Behavioral Sciences, University of Central Arkansas, AR, United States

Purpose: When assessing phonemic awareness (PA) ability, tasks that require verbal (oral) responses are often used. If verbal (oral) answers contain speech sound errors, it may be difficult for the evaluator to determine accuracy of PA skill. The present study aimed to investigate the relationship between speech sound production and PA in first and second grade children with and without speech sound disorders (SSD) using a nonverbal (non-oral) response format.

Methods: Thirty children with SSD (M age=7;3; SD =0.6) and 30 children with typical development (TD) (M age=7;3; SD =0.6) were given measures of speech sound production and phonemic awareness. Two measures of PA were administered: (1) a static assessment of phoneme deletion that required a verbal (oral) response and (2) a dynamic assessment of phoneme deletion that did not require a verbal (oral) response. Results from the two PA tasks and a speech sound production measure were compared.

Results: Results showed that performance on measures of speech sound production and PA were significantly correlated ($r=0.59$) and a strong correlation between a verbal (oral) measure of phoneme deletion and a nonverbal (non-oral) measure of phoneme deletion using a dynamic format ($r=-0.75$). Children with SSD scored more poorly than did the TD children on PA skills ($p<0.01$).

Conclusions: This study strengthens the findings of other researchers who found significant correlations between speech sound production and PA and that children with SSD have more difficulties with PA skills than TD children.

Keywords: Speech sound disorders, Phonemic awareness, Assessment, Dynamic assessment, Verbal responses, Nonverbal responses



Received: April 27, 2018
Accepted: August 22, 2018

Correspondence:
Donna Thomas

Department of Health and Human Sciences, White Hall 206, 310 West Dakota Street, PO Box 10863, Hammond, LA 70402, United States
Tel: +985-974-1003
Fax: +985-549-5030
E-mail: donna.thomas@selu.edu

INTRODUCTION

High stakes testing and performance-based funding models increase the pressure to accurately identify reading problems in children as early as possible. Assessments used to measure reading-related skills, such as phonemic awareness (PA) have conventionally require verbal (oral) responses. Consequently, measures that require verbal (oral) responses may not accurately measure skills (e.g., oral reading fluency or word finding) in children who exhibit communication issues, specifically, children who do not speak clearly and fluently [1,2], such as children with speech sound disorders (SSD).

© 2018 The Korean Association of Speech-Language Pathologists

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Children with SSD have been described in the literature as a population at risk for not developing the necessary PA skills for reading, and accordingly, acquiring competent literacy skills [3,4]. An SSD has been defined as a “significant delay in the acquisition of articulate speech sounds” [5, p. 1294]. Fifty to 75% of children found to have SSD in preschool were found to have academic problems when they were reached school-age [6,7]. Raitano et al. [8] found that children with SSD aged five to six years performed more poorly than did those with typically developing phonological systems. Carroll et al. [9] suggested that some children have difficulties in learning to read as a result of underlying problems in the use of stored phonological information. They proposed that children with SSD have problems with the sounds of the language at the level of the underlying phonological representation that will affect the ability to adequately access the sound symbol association required for decoding novel words.

However, children with SSD are a varied population [10] and do not all exhibit problems in learning to read [11]. In a study completed by Snowling et al. [12], the authors found no differences in reading and spelling skills when comparing children who had speech problems in preschool reassessed at age 15 when compared to age-matched peers. Consequently, examining the relationship between SSD and PA in young children has the potential to clarify the ability to process phonological information as it relates to speech and later reading development.

Therefore, the purpose of this study was to investigate the relationship between SSD and PA in children in the first and second grade. Further, the performance of two groups of children, those with SSD and typically developing (TD) children without SSD were compared on two measures of PA (i.e., phoneme deletion); one given in a static format that required a verbal (oral) response and one given in a dynamic assessment format that did not require a verbal (oral) response. The static test was a typical phoneme deletion task used to assess PA where the child was asked to say a word that was left over after a sound was deleted (see Appendix A), and a dynamic test that included a graduated prompting system and pictures and the child was asked to point to a picture of the word that was left over after a sound was deleted (see Appendices B and C). The dynamic test was used to assess the child’s ability to exhibit learning and transfer from the graduated prompting system as well as respond without using speech.

Phonological awareness and phonemic awareness

Since the 1980s, there has been a considerable amount of re-

search in the area of phonological awareness and PA and its relationship to early reading acquisition [13-16]. The terms phonological awareness and PA are often used to represent the same construct; however, children demonstrate this skill in different ways. While both phonological awareness and PA refer to the understanding of larger units of spoken speech (i.e., words, syllables, onsets and rimes), PA refers to the more specific ability to understand and manipulate the smaller units of spoken speech (i.e., phonemes) [17]. Both phonological awareness and PA are necessary in developing reading, development of PA of its relationship to later reading skill has received more attention [13].

Ehri et al. [18] suggests that when children understand PA, they are able to manipulate phonemes by isolating, blending, segmenting, and deleting one for another. Adequate PA skill has been shown to be consistently and highly predictive of success in reading and spelling [19]. Consequently, the success that a child has in learning to read is supported by the level of PA he/she has when entering school [13]. Thus, TD children will be able to complete PA (i.e., phoneme deletion) by 8 years of age [20]. If children have problems with developing expressive phonology, they may not develop necessary PA skills because of the difficulty accessing the sound-symbol association required for word reading [9].

Phonemic awareness assessment

Most children are typically assessed on PA during their first few years of formal schooling (i.e., kindergarten, first, and second grades) using informal assessments or standardized, norm-referenced assessments that require verbal (oral) responses. For example, in his seminal work, Bruce [20] assessed PA using an informal assessment in TD 5-, 6, and 7-year olds by asking them to say words that were left over after a sound was deleted. Many children in elementary school are assessed on PA skills using a standardized, norm-referenced assessment such as the *Comprehensive Test of Phonological Processing-2* (CTOPP-2) [21]. The CTOPP-2 is designed to measure PA skills by requiring verbal (oral) responses to phoneme blending, deletion, and elision tasks. This may be problematic for assessors when scoring verbal (oral) responses from children who exhibit speech sound errors, such as children with SSD. Guidelines for scoring verbal (oral) responses that contain speech sound errors in the studies and manuals reported here are not provided nor have guidelines in general been included for testing learners who have other expressive speech and language difficulties.

In addition, the aforementioned PA assessments (i.e., Bruce's (1964) informal PA deletion task and the PA tasks assessed on the CTOPP-2) are considered to be static, one-point in time measures. The test items are presented one at a time and the child responds to each one without feedback or assistance from the examiner [22]. Thus, data from a static assessment reflects a score or set of scores of a child's previous learning (i.e., knowledge that has been acquired up to this point in time) and the results are used for evaluation, identification, and classification purposes.

In contrast, there are a number of paradigms that have been used to assess PA via a nonverbal (non-oral) recognition type format administered as a static, one-point in time measure. For PA tasks that do not require a response via speech, children are asked to reflect on their own understanding of the composition of a word to represent their foundational ability with regards to PA. For example, Lance et al. [23] assessed implicit PA by asking TD 6-, 7-, and 8-year olds to judge words that had acceptable consonant clusters using in nonsense word-pairs by choosing the first or second word in each pair. In addition, The *Preschool and Primary Inventory of Phonological Awareness* (PIPA) [24] designed to measure PA skills in persons aged 3-6:11 years old, also requires nonverbal (non-oral) responses to syllable segmentation skills (i.e., another type of PA task).

Other researchers have examined how well children performed on assessments of phoneme deletion that used recognition or non-oral response format. For example, Claessen et al. [25] measured underlying PA ability by asking 7- and 8-year-old TD children to point to a picture that represented a word after a sound was deleted. The authors concluded that their measure, silent deletion of phonemes task (SDP), was acceptable to use for PA assessment with children who exhibit expressive speech/language difficulties because of its recognition response type format. Furthermore, other authors have previously raised concern about the use of oral response test formats for children with word finding and fluency difficulties specifically in the areas of measuring reading fluency and determining reading instruction levels [1,2]. Again, supporting the notion that nonverbal (non-oral) assessment for such skills may provide greater insight to the underlying mental processing system.

Dynamic assessment

In an extension of Claessen and colleagues' [25] work, Gillam et al. [26] assessed phoneme deletion ability with a similar

paradigm that included a dynamic component. A dynamic assessment has been used as an alternative or supplemental approach to static measures. Whereas static assessment offers no scaffolding or social support for learning, dynamic assessment embodies both the process and product of the child's learning [27]. In a dynamic assessment approach, the child is given scaffolding, social support for learning, and intervention when they fail. Dynamic assessment provides both diagnostic and prescriptive information [28] that can be applied to a wide variety of behaviors as a nonbiased way to measure language skills [22].

Dynamic assessment, therefore, is used to examine the response of the learner during the process or task and shows potential for learning, thus, can provide diagnosis and basis for treatment [29]. One method of dynamic assessment, graduated prompting, allows for cues to be given that assist a learner as they perform a task and are given to support learning and transfer [27]. For instance, if a child has trouble with phoneme deletion tasks, an examiner should engage the child in tasks of phoneme deletion during the dynamic assessment. The prompts that are developed as a hierarchy of predetermined cues are provided to give varying levels of support throughout the task, thus allowing for learning and expected transfer to occur.

Accordingly, Gillam and colleagues [26] developed a phoneme deletion task using a dynamic assessment that did not require responses via speech. Their task incorporated a prompting system that was based on response errors, which provided insight into the types of errors made and how well each participant responded to the prompts. They advocated this assessment as a way to measure PA ability in children whose speech output is distorted, such as those with SSD. For this task, the children were asked to point to pictures rather than respond verbally (orally). The authors compared the performance of 64 TD children (mean age of 8;0) on a nonverbal (non-oral) dynamic assessment of phoneme deletion (NDAPD) task to the same children's performance on a traditional static phoneme deletion (SPD) task using the same test items presented without pictures that required verbal (oral) responses. Their results indicated high correlations ($r=0.84$) between the dynamic and static tasks for the TD children, as would be expected. In addition, acceptable levels of internal consistency reliability were demonstrated ($\alpha=0.88$) between the dynamic and static tasks, suggesting that these tasks measured a similar construct. The authors suggest that the use of pointing versus verbal (oral) responses may be a viable re-

sponse mode to examine phoneme deletion ability in children whose speech output is unclear, such as children with SSD.

Aims and research questions

This study aimed to investigate the relationship between SSD and PA in children in the first and second grade. Further, the performance of two groups of children, those with SSD and TD children without SSD were compared on two measures of PA (i.e., phoneme deletion); one given in a static format that required a verbal (oral) response and one given in a dynamic assessment format that did not require a verbal (oral response). Children with SSDs have compromised verbal (oral) output, which affects the accuracy of verbal (oral) responses [30]. A nonverbal (non-oral) measure may allow the children with SSDs to complete phoneme deletion tasks by pointing to pictures in an attempt to assess phonological representation more accurately. A dynamic assessment with a graduated prompting system was used to determine whether this type of approach differentiated children with SSDs from their TD peers. A static measure of phoneme deletion was included for which verbal (oral) responses were required. It was through these measures that the explicit PA abilities of children with SSDs were assessed. The specific research questions under investigation are:

1. Is there a significant correlation between speech sound production and PA performance among learners (i.e., SSD vs. TD)?
2. Is there a significant difference in the performance of learners with and without SSD on oral response or non-oral response PA tasks?
3. Does SSD severity impact PA skills?

METHODS

Participants

Participants were recruited from three elementary schools in a small school district in Louisiana. School performance scores for these three schools (approximately 1,400 students in grades kindergarten through five) showed 25-36% below basic performance and 12% of these three schools population received special education services (which includes those receiving speech-language therapy services). Teachers of first and second graders were asked to refer potential participants who met the following inclusion criteria: were in the first or second grade, had normal auditory acuity, had normal or corrected vision as per school nurse and/or parent report; had

no identified developmental delays or emotional disorders; and were monolingual English speakers. In addition, none of the participants were receiving additional services such as special reading instruction as per teacher report. Speech-language pathologists (SLPs) who provided services to first and second graders for speech impairment and who also met the above inclusion criteria were asked to refer participants to the study.

Group selection

As participants entered the study, the *Primary Test of Non-verbal Intelligence* (PTONI) [31] and the *Peabody Picture Vocabulary Test-Fourth Edition* (PPVT-4) [32] were administered. All participants scored within normal limits (equal to or greater than a standard score of 85) on both of these measures.

Participants were also administered the sounds-in-words subtest of the *Goldman-Fristoe Test of Articulation-2* (GFTA-2) [33] by the first author to assess production of speech sounds in target words and were assigned to one of two groups (i.e., SSD group and TD group), based on the number of errors on the GFTA-2. The GFTA-2 results are reported as linear standard scores of a normal distribution with a mean of 100 and a standard deviation of 15. However, these scores cannot be interpreted in terms of the normal distribution given that articulation ability is not normally distributed. Most children after a certain age are “normal” and make one or zero errors on the GFTA-2, thus, skewing this normal distribution. For example, one participant in this sample, a female aged 8:3, exhibited four errors, which resulted in standard score of 96 and a percentile of 16, indicating at this age, 84% of the standardization sample made no errors and that only 16% of the sample made that many errors or more. According to the GFTA-2 norms, the percentile of girls this age exhibiting no errors is approximately at >48, indicating that 48% make one or more errors and 52% make no errors. In addition, dialectal sensitive scoring was taken into account in this population of participants on the GFTA-2 responses.

Additionally, percent consonants correct (PCC) adapted from Shriberg et al. [34] was calculated using the single-word responses obtained from the administration of the GFTA-2. For this study, the adapted PCC was calculated by dividing the number of correct consonants the child produced by the total number of proposed consonants (77) and multiplying by 100. Although PCC is typically calculated using running speech, this calculation was used in this study as an indicator of sever-

ity. In order to include participants in the SSD group who continue to exhibit speech sound errors that are not typical for their age, participants who exhibited more than two errors (below 96% PCC) were placed in the SSD group and those with two or fewer errors (above 96%) were placed in the TD group. The final sample was comprised of 60 participants; 30 in the SSD group ($M=82.6$; $SD=10.9$) and 30 in the TD group ($M=108.1$; $SD=6.1$). Table 1 contains the demographics of participants per group. Table 2 contains the means and standard deviations for ages and scores by inclusion criteria.

Approval to conduct this study was granted by the Institutional Review Board University of Central Arkansas Sponsored Programs Office; informed parental consent and participant assent were obtained for all participants.

General procedures

Testing was conducted in a small, quiet room on the school premises for 59 of the participants over two, one-hour sessions on separate days to minimize the amount of time that the participant was out of the classroom. One participant was assessed at a university speech and hearing clinic with a short break between the two sessions. Testing of all participants occurred within a six-week period of time.

Table 1. Demographics of participants per group

	SSD Group n=30		TD group n=30	
	Grade 1	Grade 2	Grade 1	Grade 2
AA female	4	0	9	0
AA male	8	1	1	2
C female	6	4	7	3
C male	3	4	1	7

AA, African American; C, Caucasian.

Table 2. Means, Standard deviations, and Minimum and Maximum Scores by Inclusion Criteria Group

	SSD group n=30				TD group n=30			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Age	7;3	0;6	6;4	8;4	7;3	0;6	6;4	8;3
PTONI	102.8	12.3	87	131	102.3	14.3	85	133
PPVT-4	98.7	12.6	85	131	108.3	12.7	87	135
GFTA-2	82.6	10.9	52	96	108.1	6.1	104	139
PCC	83.7	7.5	62	95	99.8	0.8	97	100

Age (years; months); PTONI, Primary Test of Nonverbal Intelligence (Ehrler & McGhee, 2008); PPVT-IV, Peabody Picture Vocabulary Test-Fourth Edition (Dunn & Dunn, 2007); GFTA-2, Goldman-Fristoe Test of Articulation-2 (Goldman & Fristoe, 2000) (PTONI, PPVT-IV, GFTA-2 standard scores reported); PCC, Percent Consonants Correct (Shriberg & Kwiatkowski, 1982).

Descriptive measures

Nine trained speech-language pathology graduate students administered the pure-tone audiometric screening, the PPVT-4, and the PTONI in alternating order as the participants entered the study under the supervision of the first author. All researchers completed a two-hour training session on the administration and scoring of the assessments conducted by the first author. In addition, all researchers practiced administration and scoring of the measures on each other under the supervision of the first author and feedback and clarification was provided.

PA tasks

The first author administered two measures of phoneme deletion, NDAPD, and SPD, both developed by Gillam et al. [26] to the participants. Both tasks required processing of stimuli presented verbally (orally), however, the NDAPD did not require a response via speech in contrast to the SPD, which required a verbal (oral) response.

NDAPD task

The participants were asked to respond by pointing to 21 sets of pictures (four pictures per card) that represented the 21 test items. The items were arranged in increasing syllable complexity, including deletion of initial and final consonant singletons and clusters. Each of the 21 sets of pictures consisted of one correct response and three foils. For example, the stimulus item contained four pictures: “car, yard, card, park” shown in line on one card (stimuli = card without /d/, target = car) (Appendix B).

Prior to the administration, each participant looked at each set of stimulus items (four pictures per card) one at a time. The researcher named each picture on the card for all sets of

pictures, in random order, and asked the participant to point to each picture as it was named. If the participant did not point to the correct picture after it was named and pointed to 5 times by the researcher, the test was discontinued and the child was excluded from the study. No participants were excluded because they did not point to the correct picture named by the researcher.

After all of the pictures were presented and comprehended, an example, two practice items, and the 21 stimulus items were administered (Refer to Appendix C). The systematic prompting and scoring system developed by Gillam et al. [26] was used to provide feedback if the participant responded incorrectly. The scores provided to the participant ranged from 1 (responded after no prompt) to 5 (responded after maximum prompting). The total possible score ranged from 21 (all correct) to 105 (maximum prompts needed). General instructions given for each item were as follows: *"Listen carefully; I am going to say a word and leave off a sound. When I leave off the sound, a real word will be left over. The real word that is left over is one of the words on the card (pointing to pictures on the stimulus card). Ready?"* For stimulus number 1, the researcher said the first prompt, *"mice without /m/ is _____."* If the participant pointed to the correct picture after hearing the first item, this was scored 1 and no further prompts were given. If the participant pointed to a picture other than the correct response, or foil type, specific prompts were given as follows:

- Foil type 1, Same picture as the stimulus; the researcher said, *"No, (stimulus; mice) without /m/ isn't (same; mice). Mice and mice are the same word."* The researcher pointed to each picture as they were named and said, *"Listen (mice - mice). I want you to show me which picture is (stimulus; mice without [sound] /m/)."*
- Foil type 2, Similar picture as the stimulus; the researcher said, *"No, (stimulus; mice) without /m/ isn't (similar; dice). Mice and dice end with the same sound."* The researcher pointed to each picture as they were named and said, *"Listen (mice /sssss/, dice /sssss/ highlighting the part of the similar word that is the same as in the stimulus word). I want you to show me which picture is (stimulus; mice without [sound] /m/)."*
- Foil type 3, Dissimilar; the researcher said, *"No, (stimulus; mice) without /m/ isn't (dissimilar; eye). Mice and eye have the same middle sound."* The researcher pointed to each picture as they were named and said, *"Listen (mice /aI/, eye /aI/). I want you to show me which picture is*

(stimulus; mice without [sound] /m/)." If the participant chose the same incorrect answer more than once, the foil type response was given.

The prompting system corresponded to a scoring scale that ranged from 1 to 5. This system allowed for the documentation of the number of prompts the participant needed before the correct answer was obtained. A participant was provided prompts based on their response to each stimulus. A participant could receive up to 5 prompts, which included hearing the first item, for each stimulus item if needed.

The following scoring system was used:

- Score of 1 = correct answer after hearing the first item and no further prompt was given.
- Score of 2 = correct response after 1 foil type prompt.
- Score of 3 = correct response after 2 foil type prompts.
- Score of 4 = correct response after 3 foil type prompts.
- Score of 5 = correct response after 4 foil type prompts.

SPD task

The SPD (Appendix A) [26] was administered to the participants. The SPD utilized the same 21 test items in the NDAPD, however, the test items were presented verbally (orally) without pictures and a verbal (oral) response was required. For example, the participant was asked to tell the researcher what the word was after deleting a sound, such as *"say the word mice without /m/."* This task consisted of 21 one-syllable words that measured correct verbal (oral) production of the target response. One training item and two practice items with corrective feedback were given prior to the stimulus items. Corrective feedback was not given during administration of the 21 stimulus items. All 21 stimulus items were administered. A correct answer was scored a 1 if the participant responded with the correctly articulated word. An incorrect answer that did not mimic the correctly articulated word was scored a 0. General instructions given for each item were as follows: *"Listen carefully; I am going to say a word and leave off a sound. When I leave off the sound, a real word will be left over. When I leave off the sound, a real word will be left over. Tell me the real word that is left over. Ready?"* For stimulus number 1, the researcher said the first prompt, *"mice without /m/ is _____."*

Data scoring

Prior to administration of the PA tasks, the participants were assigned a number from 1 to 60. To control for order effects, the odd numbered participants were administered the SPD

first and the even numbered participants were administered the NDAPD first. All of the sessions were audio recorded using a Sony ICD-PX312 Digital Flash Voice Recorder. Each standardized assessment was administered according to the standard protocol explained in the test manual. The participants' verbal (oral) responses of the GFTA-2 and the SPD were broadly phonetically transcribed using the phonetic symbols of the International Phonetic Alphabet (IPA). Because the PPVT-4, PTONI, and NDAPD required a nonverbal (non-oral) response, answers were recorded on the protocols by the researcher after the participant responded to each stimulus item.

Scoring fidelity

The first author and a licensed SLP blind to the study listened to a random 10% of the audiotapes of the verbal (oral) responses on the GFTA-2 and SPD. Inter-judge reliability was 99% and 100% respectively. Intra-judge reliability assessed by the first author on two separate occasions was 100% for both measures. The first author and a graduate student rescored 10% of the PPVT-4, PTONI, NDAPD protocols. Point-to-point percentage agreement for the number of correct productions and number of errors was calculated with the following formula: [total agreements/(total agreements+total disagreements)]. Inter-judge reliability was high for the PPVT-4 (95%), PTONI (100%), and NDAPD (100%). Intra-judge reliability was high for the PPVT-4 (100%), PTONI (98%), and NDAPD (100%).

RESULTS

The results are presented below according to the three questions posed in this investigation. Statistical comparisons and a summary of results can be found in Tables 3, 4, 5, and 6.

Relationship between speech sound production and PA proficiency

To answer the first question, "Is there a significant correlation between speech sound production and PA performance among

learners (i.e., SSD vs. TD)?" Pearson Product-Moment Correlation Coefficients were calculated. The raw scores (number of errors) from the GFTA-2 and PCC were compared with the raw scores on the NDAPD and SPD. The raw scores on the NDAPD ranged on scale from 21 through 105 and on the SPD from 0 through 21. Thus, a high score on the NDAPD indicated weak performance on PA skills, whereas, a high score on the SPD indicated strong PA skills. There was a moderate, positive significant correlation ($n=60$) between the number of speech sound errors (GFTA-2 raw score) and performance on the NDAPD ($r=0.59$), suggesting participants with higher raw scores (more speech sound errors) had higher scores on the NDAPD. There was a moderate, positive correlation between PCC and the SPD ($r=0.58$), suggesting participants with a lower PCC (more speech sound errors) had lower scores on the SPD. The participants who had more difficulty on the measures of speech sound production also had more difficulty with both phoneme deletion tasks. Additionally, there was a moderate, negative correlation between the GFTA-2 and the SPD ($r=-0.57$), suggesting the more speech sound errors a participant exhibited, poorer verbal (oral) PA skills were demonstrated.

Differences between participants with and without SSD

To answer the second question, "Is there a significant difference in the performance of learners with and without SSDs on oral response or non-oral response PA tasks?", a compara-

Table 4. Pearson Product-Moment correlation coefficients

Measure	GFTA-2	PCC	NDAPD	SPD
GFTA-2	-	0.99**	0.59**	-0.57**
PCC		-	0.59**	0.58**
NDAPD			-	-0.75**
SPD				-

** $p < 0.01$

GFTA-2, Goldman-Fristoe Test of Articulation-2 (Goldman & Fristoe, 2000); PCC, Percent Consonants Correct (Shriberg & Kwiatkowski, 1982); NDAPD, Nonverbal dynamic assessment of phoneme deletion (Gillam et al., 2011); SPD, Static phoneme deletion task.

Table 3. Means, Standard deviations, and Minimum and Maximum Scores for the Assessments by Group

	SSD group n=30				TD group n=30			
	Mean	SD	Min	Max	Mean	SD	Min	Max
NDAPD	84.5	9.4	65	99	90.6	17	8	100
SPD	41.0	24.0	0	90	73.2	24.4	19	100

NDAPD (Nonverbal (non-oral) dynamic assessment of phoneme deletion task; Gillam et al., 2011) raw score (total possible=2-105) and SPD (Static phoneme deletion task) raw score (total possible=0-21) were computed into percentages.

Table 5. Means and Standard Deviations for the by SSD Group by Severity

	Group 1 n=30		Group 2 n=17		Group 3 n=13	
	Mean	SD	Mean	SD	Mean	SD
GFTA-2	0.13	0.51	8.24	1.71	18.38	3.97
PCC	99.80	0.76	89.29	2.28	76.37	9.67
NDAPD	26.40	5.72	31.88	7.05	37.38	7.50
SPD	15.33	5.72	9.41	5.79	7.38	3.71

GFTA-2, goldman-fristoe test of articulation-2 (Goldman & Fristoe, 2000); PCC, percent consonants correct (Shriberg & Kwiatkowski, 1982); NDAPD, nonverbal dynamic assessment of phoneme deletion task (Gillam et al., 2011); SPD, static phoneme deletion task (Gillam et al., 2011) (Mean raw scores reported for GFTA-2, NDAPD, and SPD).

Table 6. ANOVA Summary Table

		Sum of squares	df	Mean square	F	Sig.
SPD	Between groups	723.872	2	361.936	13.960	0.000*
	Within groups	1,477.861	57	25.927		
	Total	2,201.733	59			
NDAPD	Between groups	1,151.292	2	575.646	13.558	0.000*
	Within groups	2,420.042	57	42.457		
	Total	3,571.333	59			

* $p < 0.05$.

SPD, static phoneme deletion task (Gillam et al., 2011); NDAPD, nonverbal dynamic assessment of phoneme deletion task (Gillam et al., 2011).

tive analysis was conducted using a two-way ANOVA, groups by conditions; two groups (SSD and TD) by two conditions (NDAPD and SPD), to determine significance and interactions. The participants' raw scores on the NDAPD and the SPD were transformed into percentages and served as the dependent variable. Partial eta squared is reported, of which .01 is small, 0.06 is medium, and 0.14 is large, to characterize the magnitude of effects.

A significant main effect was shown for group [$F(1, 58) = 20.54, p < 0.01$, partial $\eta^2 = 0.26$] and test [$F(1, 58) = 111.02, p < 0.01$, partial $\eta^2 = 0.66$]. The participants who exhibited SSD performed more poorly than did the participants who were typically developing on both tasks. A significant interaction was shown for group and test [$F(1, 58) = 21.05, p < 0.01$, partial $\eta^2 = 0.27$]. There were differences between the NDAPD and SPD across the participants with SSD and those without. A significant main effect was shown for the TD and SSD groups on the SPD [$F(1, 58) = 26.68, p < 0.01$] and the NDAPD [$F(1, 58) = 20.37, p < 0.01$]. The participants in the SSD scored more poorly than did the participants in the TD group on both the SPD and the NDAPD.

Differences based on the severity of SSD

To answer the third question, "Does SSD severity impact PA skills?," a two-way ANOVA was conducted, groups by conditions; three groups (severity) by two conditions (NDAPD and SPD) to determine significance and interactions. The PCC was divided by the following degrees of severity: 1) group 1 (> 2 errors) had a score of 97% or above; 2) group 2 (mild) had a score between 96% and 85%; and 3) group 3 (moderate) had a score below 85%. Table 5 reflects mean raw scores for the GFTA-2, SPD and NDAPD, and PCC for the three groups.

The results of the ANOVA were significant for the SPD [$F(2, 59) = 13.96; p = 0.000$] and the NDAPD [$F(2, 59) = 13.56; p = 0.000$]. There were significant differences between group 1 (TD) and group 2 (mild) and between group 1 (TD) and group 3 (moderate) on the NDAPD ($p < 0.05$). There were significant differences between group 1 (TD) and group 2 (mild) and between group 1 (TD) and group 3 (moderate) on the SPD ($p < 0.05$). However, there were no significant differences between the two groups of participants with SSD (groups 2 and 3) on the NDAPD or the SPD ($p > 0.05$). Table 6 shows the results of the ANOVA.

DISCUSSION

This study was designed to explore the relationship between speech sound production and PA tasks in children with and without SSD. Answers to questions posed in this study contribute to the understanding of the nature of SSD and its relationship to PA and the use of a dynamic assessment using a graduated prompting system (NDAPD) that did not require a verbal (oral) response. Given that the most common tasks used to measure PA require a verbal (oral) response, it makes the evidence difficult to apply to children with SSD [4]. The NDAPD was included in this study because it (1) did not require a verbal (oral) response from the children and (2) in-

cluded an error-specific graduated prompting system. Graduated prompting systems used in dynamic assessment allow for children to receive incremental levels of support to perform a task [27] and thus, encourage a child's modifiability, or ability to change given this form of support [35,36].

Relationship between speech sound production and pa proficiency

The results indicated a moderate, positive significant correlation between SSD and PA ($r=0.59$). The children who exhibited speech sound errors had more difficulty completing PA tasks. Additionally, a strong correlation was also found between the two PA measures (i.e., NDAPD and SPD) ($r=-0.75$), suggesting that these measures assess PA proficiency, regardless of whether a verbal (oral) response or nonverbal (non-oral) response was required. These results are similar to those found by Claessen et al. [25] and Gillam et al. [26] who reported high correlations between PA measures that do or do not require a verbal (oral) response.

Differences between participants with and without SSD

The results indicated a difference between the performance of children with and without SSD on PA tasks. The children who exhibited SSD scored more poorly than did the TD children as evidenced by a significant main effect ($p<0.01$). This extends the work of other researchers who also found that children who have SSD show difficulties with PA [4,37,38]. The children with SSD, even with the error-specific prompting that was provided on the NDAPD, had more difficulty with performing phoneme deletion tasks, suggesting that these children with SSD may have not fully developed underlying phonological representations to perform higher-level PA tasks [38].

Differences based on the severity of SSD

There were not significant differences between the children with mild SSD and moderate SSD ($p>0.05$). Although there was no attempt in this study to determine causes of the SSD in the participants, this particular component has clinical importance for early identification and intervention for preschool and early elementary children. It also reinforces the notion that children with non-developmental speech sound errors need preventative PA intervention, as all of the participants in this study with SSD, regardless of the severity, had difficulty demonstrating PA skills. This result is consistent with Cowan and Moran's work (1997) suggesting that even chil-

dren with mild SSD are at risk for PA development. Further, the children with SSD performed more poorly on the SPD than the NDAPD.

Clinical Implications

The findings from this study of participants with SSD and their PA proficiency have a number of clinical implications. First, the NDAPD is a relevant and valid tool to investigate phonological representations in children who are unable to produce accurate verbal (oral) speech sound production, such as those with SSD. This study reinforces the earlier findings by Gillam et al. [26] and demonstrates the NDAPD as a measure to use for phoneme deletion providing an alternative PA assessment for individuals for whom verbal (oral) measures are contraindicated. Additionally, a strong, negative correlation was also found between the two measures of phoneme deletion ($r=-0.75$) (i.e., as scores increase on the NDAPD, scores decrease on the SPD). While static measures of phoneme deletion have been demonstrated to be valid measures of PA [16, 20,23,36], the strong correlation between the two measures suggests the NDAPD as an acceptable measure of PA proficiency.

Second, using both the NDAPD and the SPD in combination could provide insight into a child's input processing skills and phonological representations. For example, a lower score on the NDAPD would reflect that the child has adequate underlying phonological representations and does understand how to complete higher-level PA tasks. On the other hand, poor performance on both the SPD and NDAPD would reflect poorly specified phonological representations that have an effect on the ability to complete higher-level PA tasks. This information can be used to develop the focus of interventions to include the development of specific phonological representations, which aids in the development of decoding ability.

Third, the NDAPD is available for SLPs to use as a measurement of PA for children who have difficulty responding verbally (orally). Verbal (oral) measures of PA do not provide a way to score responses when they differ from the norm. The use of the NDAPD would be beneficial with children who are speech that is difficult to understand or those who are non-speaking (e.g., childhood apraxia of speech or complex communication needs).

Finally, it is important to gather accurate information about children's PA ability because of the strong correlation that exists between PA and reading ability. This can be accomplished by assessing PA skills in children who exhibit SSD in order to

consider their potential for deficits in PA, which would impact the development of their reading skills. The NDAPD may be more useful than a traditional, static assessment of PA because it can provide both diagnostic and prescriptive information by analyzing the kinds of errors that the children are making. This information can then be used to design and provide intervention in this area.

Limitations and future research

First, the current study did not include a reading measure. Previous research has shown a strong correlation between PA and reading ability [16,26,38]. Future research should include an assessment of decoding ability in children with SSDs and might determine correlations between phoneme deletion assessed on the NDAPD and reading in children with SSDs.

Secondly, future investigations should consider using the NDAPD to assess PA skills in other children who have difficulty producing verbal (oral) speech, such as children who are non-speaking. Children who are non-speaking have difficulty developing age-appropriate reading skills [39]. Given that poor PA skills have been linked to poor reading skills, information about underlying phonological representations in children who are non-speaking can guide instruction in PA and reading.

Thirdly, the children in this study exhibited intact receptive vocabulary skills as measured by the PPVT-4. No additional language assessment was completed. Moreover, a measure of reading was also not included in this investigation. Future investigators should consider an assessment of all areas of language to determine if children with SSD have accompanying language and/or reading impairment. This could further quantify the effects of SSD on phonological deletion proficiency.

CONCLUSIONS

This study has made an important contribution by being the first to examine phoneme deletion proficiency in a large group of children with SSD. The results are promising, suggesting that this type of assessment is beneficial for assessing the underlying phonological representations in children who are unable to provide an accurate verbal (oral) response. The NDAPD provided additional information about what type of errors the children exhibited which can help to develop appropriate interventions for improvement in PA, which in turn, can affect later reading skills.

ACKNOWLEDGMENTS

The research was in part funded by the Office of Special Education Programs Federal Program, University of Central Arkansas Communication Sciences and Disorders Leadership Education Grant # H325D090058.

REFERENCES

1. German DJ, Newman RS. Oral reading skills of children with oral language (word-finding) difficulties. *Read Psychol.* 2007;28:397-442.
2. Reeves N, Scott K, Krieger K, Games D, Howland K, Young-Campbell L, et al. Quick: talk fast and don't stutter: the perils of oral-reading fluency assessments for children who stutter led a group of SLPs to investigate the issue and call on colleagues to change their school districts' policies. *ASHA Lead.* 2014;19:44-48.
3. Gillon GT. The efficacy of phonological awareness intervention for children with spoken language impairment. *Lang Speech Hear Serv Sch.* 2000;31:126-141.
4. Anthony JL, Aghara RG, Dunkelberger MJ, Anthony TI, Williams JM, Zhang Z. What factors place children with speech sound disorders at risk for reading problems? *Am J Speech Lang Pathol.* 2011;20:146-160.
5. Lewis BA, Shriberg LD, Freebairn LA, Hansen AJ, Stein CM, Taylor HG, et al. The genetic bases of speech sound disorders: evidence from spoken and written language. *J Speech Lang Hear Res.* 2006;49:1294-1312.
6. Flax JF, Realpe-Bonilla T, Hirsh LS, Brzustowicz LM, Bartlett CW, Tallal P. Specific language impairment in families: evidence for co-occurrence with reading impairments. *J Speech Lang Hear Res.* 2003;46:530-543.
7. Lewis BA, Freebairn LA, Taylor HG. Academic outcomes in children with histories of speech sound disorders. *J Commun Disord.* 2000;33:11-30.
8. Ratiano NA, Pennington BF, Tunick RA, Boada R, Shriberg LD. Pre-literacy skills of subgroups of children with speech sound disorders. *J Child Psychol Psychiatry.* 2004;45:821-835.
9. Carroll J, Snowling M. Language and phonological skills in children at high risk for reading difficulties. *J Child Psychol Psychiatry.* 2004;45:631-640.
10. Shriberg LD. Diagnostic markers for child speech-sound disorders: Introductory comments. *Clin Linguist Phon.* 2003;17:501-505.
11. Hesketh A. Early literacy achievement of children with a history of speech problems. *Int J Lang Commun Disord.* 2004;39:453-468.
12. Snowling MJ, Bishop DVM, Stothard SE. Is pre-school language impairment a risk factor for dyslexia in adolescence? *J Child Psychol Psychiatry.* 2000;41:587-600.
13. Adams MJ. *Beginning to read: thinking and learning about print.* Cambridge, MA: MIT Press, 1990.
14. Stanovich KE. Matthew effects in reading: some consequences of

- individual differences in the acquisition of literacy. *Read Res Q.* 1986;21:360-407.
15. Troia GA. Phonological processing and its influence on literacy learning. In : Stone C, Silliman E, Ehren B, Apel K, editors. *Handbook of language and literacy: development and disorders.* p. 271-301. New York, NY: Guilford Press. 2004.
 16. Yopp HK. The validity and reliability of PA tests. *Read Res Q.* 1988; 23:159-177.
 17. Chapman ML. Phonemic awareness: clarifying what we know. *Lit Teach Learn.* 2003;7:91-114.
 18. Ehri LC, Nunes SR, Stahl SA, Willows DM. Systematic phonics instruction helps students learn to read: evidence from the National Reading Panel's meta-analysis. *Rev Educ Res.* 2001;7:393-447.
 19. Troia GA. Phonological awareness intervention research: a crucial review of the research methodology. *Read Res Q.* 1999;34:28-52.
 20. Bruce D. An analysis of word sounds by young children. *Br J of Educ Psychol.* 1964;34:158-170.
 21. Wagner RK, Torgesen JK, Rashotte CA, Pearson NA. *Comprehensive test of phonological processing.* 2nd ed. Austin, TX: Pro-Ed, 2013.
 22. Sternberg RJ, Grigorenko EL. *Dynamic testing: the nature and measurement of learning potential.* New York: Cambridge University Press, 2002.
 23. Lance DM, Swanson LA, Peterson HA. A validity study of an implicit phonological awareness paradigm. *J Speech Lang Hear Res.* 1997;40:1002-1010.
 24. Dodd B, Crosbie S, McIntosh B, Teitzel T, Ozanne A. *Preschool and primary inventory of phonological awareness.* London: Pearson Education Limited, 2000.
 25. Claessen M, Leitão S, Barrett N. Investigating children's ability to reflect on stored phonological representations: the silent deletion of phonemes task. *Int J Lan Commun Disord.* 2010;45:411-423.
 26. Gillam SL, Fargo J, Foley B, Olszewski A. A nonverbal phoneme deletion task administered in a dynamic assessment format. *J Commun Disord.* 2011;44:236-245.
 27. Campione JC, Brown AL. Linking dynamic assessment with school achievement. In : Lidz CS, editor. *Dynamic assessment: an interactional approach to evaluating learning potential.* p. 82-115. New York: The Guilford Press, 1987.
 28. Gutierrez-Clellan VF, Peña E. Dynamic assessment of diverse children: a tutorial. *Lang Speech Hear Serv Sch.* 2001;32(4):212-224.
 29. Laing S, Kamhi A. Assessment of language and literacy in culturally and linguistically diverse populations. *Lang Speech Hear Serv Sch.* 2003;34:44-55.
 30. Anthony JL, Ahgara RG, Dunkelberger MJ, Anthony TI, Williams JM, Zhang Z. What factors place children with speech sound disorders at risk for reading problems? *Am J Speech Lang Path.* 2011; 20:146-160.
 31. Ehrler DJ, McGhee RL. *Primary test of nonverbal intelligence.* Austin, TX: Pro-Ed, 2008.
 32. Dunn A, Dunn A. *Peabody picture vocabulary test.* 4th ed. Circle Pines, MN: American Guidance Service, 2007.
 33. Goldman R, Fristoe M. *Goldman-Fristoe test of articulation-2nd ed.* Circle Pines, MN: American Guidance Service, 2000.
 34. Shriberg LD, Kwiatkowski J. *Phonological disorders III: a procedure for assessing severity of involvement.* *J Speech Hear Disord.* 1982;47:256-270.
 35. Coventry, WL and Byrne, B, Olsen RK. Dynamic and static assessment of phonological awareness in Preschool: a behavior-genetic study. *J Learn Disabil.* [Internet]. 2011 [cited 2018 Apr 12];44:322-329.
 36. Spector JE. Predicting progress in beginning reading: dynamic assessment of phonemic awareness. *J Educ Psychol.* 1992;84:353-363.
 37. Gernand KL, Moran MJ. Phonological awareness abilities in 6-year-old children with mild to moderate phonological impairments. *Commun Disord Q.* 2007;28:206-215.
 38. Gillam SL, Ford MB. Dynamic assessment of phonological awareness for children with speech sound disorders. *Child Lang Teach Ther.* 2012;28:297-308.
 39. Zascavage VT, Keefe CH. Students with severe speech and physical impairments: opportunity barriers to literacy. *Focus on Autism and Other Developmental Disabilities.* 2004;19:223-234.

Appendix A

Static Phonemic Deletion Measure

One example and 2 practice items will be used. Examiner: "Listen carefully; I am going to say a word and leave off a sound. When I leave off the sound, a real word will be left over. Tell me the real word that is left over. Ready? Card without /d/ is car. Now, you say the real word that is leftover after I leave off a sound. Say throw without /θ/". If correct, move to practice B. If incorrect, examiner: "Throw without /θ/ is row. Now, you try one. Say pink without /k/". Repeat corrective feedback if child is incorrect. Continue with test items. No corrective feedback is given for items 1 through 21. Give all stimulus items.

Practice A: throw /θ/ = row

Practice B: pink /k/ = pin

Scoring: 1 for correct response, 0 for incorrect response. "Say _____ without / /".

Stimulus	Correct Response	Child's Response	Score
1.	mice /m/	ice	
2.	cup /k/	up	
3.	tape /t/	ape	
4.	gate /g/	ate	
5.	deer /d/	ear	
6.	pine /n/	pie	
7.	boat /t/	bow	
8.	goat /g/	go	
9.	bone /n/	bow	
10.	bike /k/	bye	
11.	glue /l/	goo	
12.	blow /l/	bow	
13.	skip /k/	sip	
14.	grass /r/	gas	
15.	plants /l/	pants	
16.	please /l/	peas	
17.	snap /s/	nap	
18.	stop /s/	top	
19.	skis /s/	keys	
20.	belt /t/	bell	
21.	fast /s/	fat	

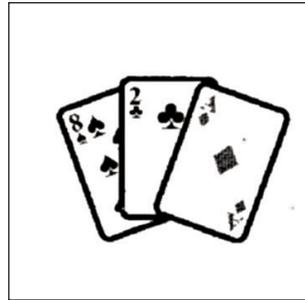
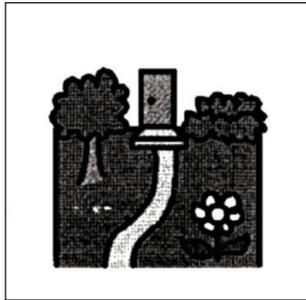
Add the number of correct responses for #1-21.

Total possible: 21

Total Score: _____

Appendix B

Example of Nonverbal, Dynamic Phoneme Deletion Picture Sets



Card without /d/ is _____

Appendix C

Nonverbal Dynamic Phonemic Deletion Task (Gillam et al., 2011)

Stimulus					Total Score
Initial Sound:					
Foil Type	Target	Similar	Same	Dissimilar	
1. mice /m/	ice	dice	mice	eye	
Foil Type	Similar	Same	Dissimilar	Target	
2. cup /k/	cap	cup	under	up	
Foil Type	Similar	Target	Dissimilar	Same	
3. tape /t/	cape	ape	eight	tape	
Foil Type	Same	Similar	Target	Dissimilar	
4. gate /g/	gate	wait	eight	ape	
Foil Type	Dissimilar	Target	Same	Similar	
5. deer /d/	eat	ear	deer	tear	
Final Sound:					
Foil Type	Dissimilar	Same	Similar	Target	
6. pine /n/	wipe	pine	vine	pie	
Foil Type	Similar	Dissimilar	Target	Same	
7. boat /t/	goat	cap	bow	boat	
Foil Type	Target	Similar	Same	Dissimilar	
8. goat (t)	go	coat	goat	bow	
Foil Type	Same	Target	Similar	Dissimilar	
9. bone /n/	bone	bow	phone	bowl	
Foil Type	Dissimilar	Same	Similar	Target	
10. bike /k/	pie	bike	hike	bye	
Initial cluster: (2nd element)					
Foil Type	Similar	Same	Dissimilar	target	
11. glue /l/	shoe	glue	boot	goo	
Foil Type	Similar	Dissimilar	Same	Target	
12. blow /l/	blue	bone	blow	bow	
Foil Type	Target	Same	Similar	Dissimilar	
13. skip /k/	sip	skip	skate	sit	
Foil Type	Dissimilar	Target	Same	Similar	
14. grass /r/	gate	gas	grass	green	
Foil Type	Dissimilar	Similar	Same	Target	
15. plants/l/	axe	ants	plants	pants	
Foil Type	Similar	Dissimilar	Same	Target	
16. please/l/	play	pink	please	peas	
Initial cluster: (1st element)					
Foil Type	Target	Dissimilar	Similar	Same	
17. snap /s/	nap	no	snow	snap	
Foil Type	Target	Same	Similar	Dissimilar	
18. stop /s/	top	stop	steak	Hop	
Foil Type	Target	Same	Similar	Dissimilar	
19. skis /s/	keys	skis	skate	Please	
Final cluster: (2nd element)					
Foil Type	Same	Dissimilar	Target	Similar	
20. belt /t/	belt	yell	bell	Melt	
Final cluster: (1st element)					
Foil Type	Similar	Same	Target	Dissimilar	
21. fast /s/	last	fast	fat	foot	
Total Score					